



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 134436

To: Sarvamangala Devi
Location: REM 3C18
Art Unit: 1645
Tuesday, October 05, 2004

Case Serial Number: 09/359426

From: Beverly Shears
Location: Remsen Bldg.
RM 1A54
Phone: 571-272-2528

beverly.shears@uspto.gov

Search Notes

Shears, Beverly

134436

From: Devi, Sarvamangala
Sent: Friday, October 01, 2004 2:18 PM
To: Shears, Beverly
Subject: 09/359,426

Beverly:

Would you please perform a sequence search for SEQ ID NO: 2 from application 09/359,426 both in commercial and interference databases?

Thanx.

S. DEVI, Ph.D.
AU 1645
Rems - 3C18



Date completed: 09-10-05-04
Searcher: Beverly, 22528
Terminal time: 21
Elapsed time:
CPU time:
Total time: 24
Number of Searches:
Number of Databases: 2

Search Site

☐ STIC
☐ CM-1
☐ Pre-S

Type of Search

☐ N.A. Sequence
☐ A.A. Sequence
☐ Structure
☐ Bibliographic

Vendors

☐ IG
☒ STN
☐ Dialog
☐ APS
☐ Geninfo
☐ SDC
☐ DARC/Questel
☒ Other CGN

Blank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 15:07:49 ; Search time 29 Seconds
(without alignments)
33.824 Million cell updates/sec

Title: US-09-359-426C-2

Perfect score: 81
Sequence: 1 XEKTPLTTAAXAPVWNA 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/6CTUS-COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	77.8	482	4	US-09-252-991A-21327
2	53	65.4	406	4	US-09-543-681A-4268
3	41	50.6	788	1	US-08-194-338-12
4	40	49.4	456	4	US-09-489-039A-13165
5	40	49.4	498	1	US-08-470-202-59
6	40	49.4	498	1	US-08-471-770-59
7	40	49.4	498	2	US-08-468-059-59
8	40	49.4	498	3	US-09-109-916-59
9	40	49.4	498	4	US-08-886-156-59
10	40	49.4	498	4	US-09-886-149-59
11	40	49.4	498	4	US-08-886-150-59
12	40	49.4	498	4	US-09-886-159-59
13	39	48.1	145	4	US-09-328-352-7139
14	37	45.7	262	4	US-09-107-532A-5791
15	37	45.7	311	4	US-09-543-681A-5777
16	37	45.7	342	4	US-09-252-991A-19222
17	37	45.7	456	1	US-08-624-125-20
18	37	45.7	456	4	US-09-937-155-20
19	37	45.7	463	2	US-08-679-635A-4
20	37	45.7	463	4	US-09-419-163-4
21	37	45.7	498	4	US-09-323-998E-57
22	37	45.7	500	4	US-09-323-998E-58
23	37	45.7	500	4	US-09-323-998E-59
24	36.5	45.1	515	4	US-09-134-000C-3767
25	36	44.4	19	3	US-08-943-173-8
26	36	44.4	58	3	US-08-943-173-16
27	36	44.4	71	3	US-08-943-173-2

28	36	44.4	162	4	US-09-252-991A-17602	Sequence 17602, A
29	36	44.4	181	3	US-09-117-257-19	Sequence 19, Appl
30	36	44.4	181	3	US-08-945-476-19	Sequence 19, Appl
31	36	44.4	181	4	US-09-489-352-19	Sequence 19, Appl
32	36	44.4	182	3	US-09-117-257-48	Sequence 48, Appl
33	36	44.4	182	4	US-09-489-352-48	Sequence 48, Appl
34	36	44.4	236	4	US-09-134-001C-3558	Sequence 358, Ap
35	36	44.4	294	4	US-09-446-301A-6	Sequence 6, Appl
36	36	44.4	295	4	US-09-099-932-6	Sequence 6, Appl
37	36	44.4	355	2	US-08-458-555-2	Sequence 2, Appl
38	36	44.4	501	4	US-09-323-998E-55	Sequence 55, Appl
39	36	44.4	507	4	US-09-446-301A-51	Sequence 51, Appl
40	36	44.4	610	1	US-07-821-717B-6	Sequence 6, Appl
41	36	44.4	610	1	US-08-119-262B-6	Sequence 6, Appl
42	36	44.4	610	1	US-08-135-929A-11	Sequence 11, Appl
43	36	44.4	610	1	US-08-234-265A-11	Sequence 11, Appl
44	36	44.4	2318	3	US-09-091-219-24	Sequence 24, Appl
45	36	44.4	2318	4	US-09-660-541-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-09-252-991A-21327
; Sequence 21327, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21327
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21327

Query Match 77.8%; Score 63; DB 4; Length 482;
Best Local Similarity 82.4%; Pred. No. 0.0089; 3; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 3

Qy 2 EEXTPLTTAAXAPVWVN 18
|||||
Db 2 EEXTPLTTAAGAPVDN 18
|||||

RESULT 2

US-09-543-681A-4268
; Sequence 4268, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILII
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4268
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4268

Query Match 55.4%; Score 53; DB 4; Length 406;
Best Local Similarity 75.0%; Pred. No. 0.35;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 3 EKTPLTTAAAXAPVXN 18
||| ||||| |||||
Db 26 EKKLTTAAGAPVVDN 41
||| ||||| |||||
RESULT 3
US-08-194-338-12
; Sequence 12, Application US/08194338
; Patent No. 5474898
; GENERAL INFORMATION:
; APPLICANT: Venter, John C.
; APPLICANT: Fraser, Claire M.
; APPLICANT: McCombie, William R.
; TITLE OF INVENTION: OCTOPAMINE RECEPTOR
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194.338
; FILING DATE: 08-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/676,174
; FILING DATE: 28-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH101.001DV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
US-08-194-338-12
Query Match 50.6%; Score 41; DB 1; Length 788;
Best Local Similarity 52.9%; Pred. No. 76;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
OY 3 EKTPLTTAAAXAPVXNA 19
||| ||||| |||||
Db 602 EDQPTTAAAPLASAA 618
||| ||||| |||||
RESULT 4
US-09-489-039A-13165
; Sequence 13165, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13165
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13165
Query Match 49.4%; Score 40; DB 4; Length 456;
Best Local Similarity 57.1%; Pred. No. 60;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 2 EKTPLTTAAAXAPV 15
||| ||||| |||||
Db 393 EQKPLTAADMAAI 406
||| ||||| |||||
RESULT 5
US-08-470-202-59
; Sequence 59, Application US/08470202
; Patent No. 5759808
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht v.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: Retrovirus from the HIV Group and Its
; TITLE OF INVENTION: Use
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470.202
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/132,653
; FILING DATE: 05-OCT-1993
; APPLICATION NUMBER: DE P 42 33 646.5
; FILING DATE: 06-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 42 35 718.7
; FILING DATE: 22-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 42 44 541.8
; FILING DATE: 30-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 43 18 186.4
; FILING DATE: 01-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Michael J. Blake
; REGISTRATION NUMBER: 37,096
; REFERENCE/DOCKET NUMBER: 05495-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-470-202-59

Query Match 49.4%; Score 40; DB 1; Length 498;
Best Local Similarity 44.4%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 EKTPLTTAAAXAPVVXNA 19
||:|:|:|:
Db 119 EETSPROTQSNYPVITNA 136

RESULT 6

US-08-471-770-59
Sequence 59, Application US/08471770
Patent No. 5770427

GENERAL INFORMATION:

APPLICANT: Guertler, Lutz G.
APPLICANT: Eberle, Josef
APPLICANT: Brunn, Albrecht v.
APPLICANT: Knapp, Stefan
APPLICANT: Hauser, Hans-Peter
TITLE OF INVENTION: Retrovirus from the HIV Group and Its
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,770
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/132,653
FILING DATE: 05-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 42 33 646.5
FILING DATE: 06-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 42 35 718.7
FILING DATE: 22-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 42 44 541.8
FILING DATE: 30-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P 43 18 186.4
FILING DATE: 01-JUN-1993
ATTORNEY/AGENT INFORMATION:

NAME: Carol P. Einaudi
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 05495-0001-03000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 59:

SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-471-770-59

Query Match 49.4%; Score 40; DB 1; Length 498;
Best Local Similarity 44.4%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 EKTPLTTAAAXAPVVXNA 19
||:|:|:|:
Db 119 EETSPROTQSNYPVITNA 136

RESULT 7

US-08-468-059-59
Sequence 59, Application US/08468059
Patent No. 5840480

GENERAL INFORMATION:

APPLICANT: Guertler, Lutz G.
APPLICANT: Eberle, Josef
APPLICANT: Brunn, Albrecht v.
APPLICANT: Knapp, Stefan
APPLICANT: Hauser, Hans-Peter
TITLE OF INVENTION: Retrovirus from the HIV Group and Its
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,059
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/132,653
FILING DATE: 05-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 42 33 646.5
FILING DATE: 06-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 42 35 718.7
FILING DATE: 22-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 42 44 541.8
FILING DATE: 30-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE P 43 18 186.4
FILING DATE: 01-JUN-1993
ATTORNEY/AGENT INFORMATION:

NAME: Carol P. Einaudi
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 05495-0001-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-468-059-59

Query Match          49.4%; Score 40; DB 2; Length 498;
Best Local Similarity 44.4%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 7; Indels 7; Gaps 0;

QY 2 EEKTLTAAAXAPVVXNA 19
DB 119 EETSPROTQSNYPVITNA 136

RESULT 8
US-09-109-916-59
; Sequence 59, Application US/09109916
; Patent No. 6277561
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/109,916
; CURRENT FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: DE P 42 33 646.5
; EARLIER FILING DATE: 1992-10-06
; EARLIER APPLICATION NUMBER: DE P 42 35 718.7
; EARLIER FILING DATE: 1992-10-22
; EARLIER APPLICATION NUMBER: DE P 42 44 541.8
; EARLIER FILING DATE: 1992-12-30
; EARLIER APPLICATION NUMBER: DE P 43 18 186.4
; EARLIER FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-109-916-59

Query Match          49.4%; Score 40; DB 3; Length 498;
Best Local Similarity 44.4%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 7; Indels 7; Gaps 0;

QY 2 EEKTLTAAAXAPVVXNA 19
DB 119 EETSPROTQSNYPVITNA 136

RESULT 9
US-09-886-156-59
; Sequence 59, Application US/09886156
; Patent No. 6528626
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,156
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-156-59

Query Match          49.4%; Score 40; DB 4; Length 498;
Best Local Similarity 44.4%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 7; Indels 7; Gaps 0;

QY 2 EEKTLTAAAXAPVVXNA 19
DB 119 EETSPROTQSNYPVITNA 136

RESULT 11
US-09-886-150-59
; Sequence 59, Application US/09886150
; Patent No. 6531587
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,150
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-149-59

Query Match          49.4%; Score 40; DB 4; Length 498;
Best Local Similarity 44.4%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 7; Indels 7; Gaps 0;

QY 2 EEKTLTAAAXAPVVXNA 19
DB 119 EETSPROTQSNYPVITNA 136

RESULT 10
US-09-886-149-59
; Sequence 59, Application US/09886149
; Patent No. 6531137
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,149
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-149-59

Query Match          49.4%; Score 40; DB 4; Length 498;
Best Local Similarity 44.4%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 7; Indels 7; Gaps 0;

QY 2 EEKTLTAAAXAPVVXNA 19
DB 119 EETSPROTQSNYPVITNA 136

RESULT 11
US-09-886-150-59
; Sequence 59, Application US/09886150
; Patent No. 6531587
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,150
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-149-59

```

; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,150
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-150-59;

Query Match 49.4%; Score 40; DB 4; Length 498;
Best Local Similarity 44.4%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 BEKPLTTAAAXAPVVXNA 19
|||:|:|:|:|:
Db 119 EETSPRTSQNYPIVTNA 136

RESULT 12

US-09-886-159-59
; Sequence 59, Application US/09886159
; Patent No. 6551824
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,159
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-159-59

Query Match 49.4%; Score 40; DB 4; Length 498;
Best Local Similarity 44.4%; Pred. No. 66;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 BEKPLTTAAAXAPVVXNA 19
|||:|:|:|:|:
Db 119 EETSPRTSQNYPIVTNA 136

RESULT 13

US-09-328-352-7139
; Sequence 7139, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7139
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7139

Query Match 48.1%; Score 39; DB 4; Length 145;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 KTELTTAAAXAPV 15
:|||||:
Db 35 RMPLTAAOAPV 46

RESULT 14

US-09-107-532A-5791
; Sequence 5791, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5791:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 262 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: (B) LOCATION 1...262
SEQUENCE DESCRIPTION: SEQ ID NO: 5791:
US-09-107-532A-5791

Query Match 45.7%; Score 37; DB 4; Length 262;
Best Local Similarity 87.5%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 EKTPLTTA 10
| | | | |
DB 12 EKTPLTTS 19

RESULT 15
US-09-543-681A-5777
; Sequence 5777, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5777
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5777

Query Match 45.7%; Score 37; DB 4; Length 311;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 EKTPLTTAAXAPV 15
| | | | |
DB 89 EKVPGTTSGVAPI 101

RESULT 16
US-09-252-991A-19222
; Sequence 19222, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19222
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19222

Query Match 45.7%; Score 37; DB 4; Length 342;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 PLTTAAXAPVXNA 19
| | | | |
DB 280 PVTPTAAPVIDKA 293

RESULT 17
US-08-624-125-20
; Sequence 20, Application US/08624125
; Patent No. 574341
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
; TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/624,125
; FILING DATE: 29-MAR-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: KELBER, STEVEN B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 2747-063-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-624-125-20

Query Match 45.7%; Score 37; DB 1; Length 456;
Best Local Similarity 72.7%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 TAAAXAPVXNA 19
| | | | |
DB 336 TLAAXAPVXNA 346

RESULT 18
US-08-937-155-20
; Sequence 20, Application US/08937155
; Patent No. 6524811
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
; TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; ADDRESS: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,155
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/624,125
FILING DATE: 29-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: KELBER, STEVEN B.
REGISTRATION NUMBER: 30,073
REFERENCE/DOCKET NUMBER: 2747-063-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 456 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-937-155-20

Query Match 45.7%; Score 37; DB 4; Length 456;
Best Local Similarity 72.7%; Pred. No. 1.9e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 TTAAXAPVYXNA 19
DB 336 TTAAXAPVYXNA 346

RESULT 19
US-08-679-635A-4
Sequence 4, Application US/08679635A
Patent No. 5985643
GENERAL INFORMATION:
APPLICANT: Tomasz, Alexander
APPLICANT: Delencastre, Herminia
TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF
TITLE OF INVENTION: METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/679,635A
FILING DATE: 10-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-141
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-679-635A-4

Query Match 45.7%; Score 37; DB 2; Length 463;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 TTAAXAPVYXNA 19
DB 382 TTAAXAPLVQTA 393

RESULT 20
US-09-419-163-4
Sequence 4, Application US/09419163
Patent No. 6391614
GENERAL INFORMATION:
APPLICANT: Tomasz, Alexander
APPLICANT: Delencastre, Herminia
TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF
TITLE OF INVENTION: METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/419,163
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/679,635
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-141
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 463 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-419-163-4

Query Match 45.7%; Score 37; DB 4; Length 463;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 TTAAXAPVYXNA 19
DB 382 TTAAXAPLVQTA 393

RESULT 21
US-09-323-998E-57
Sequence 57, Application US/09323998E

Patent No. 6642021
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998E
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Capsicum annuum
US-09-323-998E-57

Query Match 45.7%; Score 37; DB 4; Length 498;
Best Local Similarity 72.7%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 TAAXAPVVXNA 19
| | | | |
DB 370 TLAAAPVVANA 380

RESULT 22
US-09-323-998E-58
; Sequence 58, Application US/09323998E
; Patent No. 6642021
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998E
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-323-998E-58

Query Match 45.7%; Score 37; DB 4; Length 500;
Best Local Similarity 72.7%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 TAAXAPVVXNA 19
| | | | |
DB 372 TLAAAPVVANA 382

RESULT 23
US-09-323-998E-59
; Sequence 59, Application US/09323998E
; Patent No. 6642021
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019

; CURRENT APPLICATION NUMBER: US/09/323,998E
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-323-998E-59

Query Match 45.7%; Score 37; DB 4; Length 500;
Best Local Similarity 72.7%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 TAAXAPVVXNA 19
| | | | |
DB 372 TLAAAPVVANA 382

RESULT 24
US-09-134-000C-3767
; Sequence 3767, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3767
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3767

Query Match 45.1%; Score 36.5; DB 4; Length 515;
Best Local Similarity 64.3%; Pred. No. 2.7e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 5 TPLTTAAXAPVVXN 18
| | | | |
DB 379 TP-TTAATAPSIDN 391

RESULT 25
US-08-943-173-8
; Sequence 8, Application US/08943173
; Patent No. 6048538
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Shen, Fan
; APPLICANT: Chen, Pei De
; TITLE OF INVENTION: PEPTIDES DERIVED
; TITLE OF INVENTION: FROM THE NON-STRUCTURAL PROTEINS OF
; TITLE OF INVENTION: FOOT AND MOUTH DISEASE VIRUS AS
; TITLE OF INVENTION: DIAGNOSTIC REAGENTS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,173
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lin, Maria C.H.
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4152
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-943-173-8

Query Match 44.4%; Score 36; DB 3; Length 19;
Best Local Similarity 53.3%; Pred. No. 7.7;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 EKTPLTTAAAPV 16
| : ||| |
Db 5 ERQKPLKVKAKAPV 19

RESULT 26
US-08-943-173-16
; Sequence 16, Application US/08943173
; Patent No. 6048538
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Shen, Fan
; APPLICANT: Chen, Pei De
; TITLE OF INVENTION: PEPTIDES DERIVED
; TITLE OF INVENTION: FROM THE NON-STRUCTURAL PROTEINS OF
; TITLE OF INVENTION: FOOT AND MOUTH DISEASE VIRUS AS
; TITLE OF INVENTION: DIAGNOSTIC REAGENTS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,173
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lin, Maria C.H.
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4152
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-943-173-16

Query Match 44.4%; Score 36; DB 3; Length 58;
Best Local Similarity 53.3%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 EKTPLTTAAAPV 16
| : ||| |
Db 31 ERQKPLKVKAKAPV 45

RESULT 27
US-08-943-173-2
; Sequence 2, Application US/08943173
; Patent No. 6048538
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Shen, Fan
; APPLICANT: Chen, Pei De
; TITLE OF INVENTION: PEPTIDES DERIVED
; TITLE OF INVENTION: FROM THE NON-STRUCTURAL PROTEINS OF
; TITLE OF INVENTION: FOOT AND MOUTH DISEASE VIRUS AS
; TITLE OF INVENTION: DIAGNOSTIC REAGENTS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0053
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,173
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lin, Maria C.H.
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4152
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-943-173-2

Query Match 44.4%; Score 36; DB 3; Length 71;
Best Local Similarity 53.3%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 EKTPLTTAAAPV 16
| : ||| |
Db 31 ERQKPLKVKAKAPV 45

RESULT 28
US-09-252-991A-17602
; Sequence 17602, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

```
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17602
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17602

Query Match          44.4%; Score 36; DB 4; Length 162;
Best Local Similarity 63.6%; Pred. No. 87;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      4 KTLPLTTAAAXAP 14
       :||| ||| |||
Db      1 RTPAATTAIAP 11

RESULT 29
US-09-117-257-19
; Sequence 19, Application US/09117257
; Patent No. 6214355
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Guo, Betty
; APPLICANT: Hanson, Mark
; TITLE OF INVENTION: Dbpa AND Dbpb COMPOSITIONS AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/117,257
; CURRENT FILING DATE: 1998-07-22
; EARLIER APPLICATION NUMBER: PCT/US96/17081
; EARLIER FILING DATE: 1996-10-22
; EARLIER APPLICATION NUMBER: 08/589,711
; EARLIER FILING DATE: 1996-01-22
; EARLIER APPLICATION NUMBER: 08/427,023
; EARLIER FILING DATE: 1995-04-24
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 19
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-117-257-19

Query Match          44.4%; Score 36; DB 3; Length 181;
Best Local Similarity 47.1%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy      3 EKTPLTTAAAXAPVVKNA 19
       :||| ||| |||
Db      139 EKTPTTAAEGIIITAKA 155

RESULT 30
US-08-945-476-19
; Sequence 19, Application US/08945476
; Patent No. 6248517
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DECORIN BINDING PROTEIN COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF USE
; NUMBER OF SEQUENCES: 27
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,476
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/589,711
; FILING DATE: 22-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/427,023
; FILING DATE: 24-APR-1995
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-945-476-19

Query Match          44.4%; Score 36; DB 3; Length 181;
Best Local Similarity 47.1%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy      3 EKTPLTTAAAXAPVVKNA 19
       :||| ||| |||
Db      139 EKTPTTAAEGIIITAKA 155

RESULT 31
US-09-489-352-19
; Sequence 19, Application US/09489352
; Patent No. 6312807
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Guo, Betty
; APPLICANT: Hanson, Mark
; TITLE OF INVENTION: Dbpa AND Dbpb COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 4210.000500
; CURRENT APPLICATION NUMBER: US/09/489,352
; CURRENT FILING DATE: 2008-01-21
; EARLIER APPLICATION NUMBER: PCT/US96/17081
; EARLIER FILING DATE: 1996-10-22
; EARLIER APPLICATION NUMBER: 08/589,711
; EARLIER FILING DATE: 1996-01-22
; EARLIER APPLICATION NUMBER: 08/427,023
; EARLIER FILING DATE: 1995-04-24
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 19
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Borrelia afzelii
US-09-489-352-19

Query Match          44.4%; Score 36; DB 4; Length 181;
Best Local Similarity 47.1%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy      3 EKTPLTTAAAXAPVVKNA 19
       :||| ||| |||
Db      139 EKTPTTAAEGIIITAKA 155

RESULT 32
US-09-117-257-48
; Sequence 48, Application US/09117257
; Patent No. 6214355
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Guo, Betty
; APPLICANT: Hanson, Mark
; TITLE OF INVENTION: Dbpa AND Dbpb COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: 4210.000500
; CURRENT APPLICATION NUMBER: US/09/117,257
; CURRENT FILING DATE: 1998-07-22
```



```
/ EARLIER APPLICATION NUMBER: PCT/US96/17081
/ EARLIER FILING DATE: 1996-10-22
/ EARLIER APPLICATION NUMBER: 08/589,711
/ EARLIER FILING DATE: 1996-01-22
/ EARLIER APPLICATION NUMBER: 08/427,023
/ EARLIER FILING DATE: 1995-04-24
/ NUMBER OF SEQ ID NOS: 66
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 48
/ TYPE: PRT
/ LENGTH: 182
/ ORGANISM: Borrelia afzelii
US-09-117-257-48

Query Match      44.4%; Score 36; DB 3; Length 182;
Best Local Similarity 47.1%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPVXNA 19
Db 139 EKTPTTAAEGITIAKA 155

RESULT 33
US-09-489-352-48
/ Sequence 48, Application US/09489352
/ Patent No.: 6312907
/ GENERAL INFORMATION:
/ APPLICANT: Hook, Magnus
/ APPLICANT: Guo, Betty
/ APPLICANT: Hanson, Mark
/ TITLE OF INVENTION: DcpA AND DcpB COMPOSITIONS AND METHODS OF USE
/ FILE REFERENCE: 4210.000500
/ CURRENT APPLICATION NUMBER: US/09/489,352
/ CURRENT FILING DATE: 2000-01-21
/ EARLIER APPLICATION NUMBER: PCT/US96/17081
/ EARLIER FILING DATE: 1996-10-22
/ EARLIER APPLICATION NUMBER: 08/589,711
/ EARLIER FILING DATE: 1996-01-22
/ EARLIER APPLICATION NUMBER: 08/427,023
/ EARLIER FILING DATE: 1995-04-24
/ NUMBER OF SEQ ID NOS: 66
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 48
/ TYPE: PRT
/ LENGTH: 182
/ ORGANISM: Borrelia afzelii
US-09-489-352-48

Query Match      44.4%; Score 36; DB 4; Length 182;
Best Local Similarity 47.1%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPVXNA 19
Db 139 EKTPTTAAEGITIAKA 155

RESULT 34
US-09-134-001C-3558
/ Sequence 3558, Application US/09134001C
/ Patent No.: 6380370
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
/ FILE REFERENCE: GTC-007
/ CURRENT APPLICATION NUMBER: US/09/134,001C
/ CURRENT FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/064,964
/ PRIOR FILING DATE: 1997-11-08
/ PRIOR APPLICATION NUMBER: US 60/055,779
/ PRIOR FILING DATE: 1997-08-14

/ EARLIER APPLICATION NUMBER: PCT/US96/17081
/ EARLIER FILING DATE: 1996-10-22
/ EARLIER APPLICATION NUMBER: 08/589,711
/ EARLIER FILING DATE: 1996-01-22
/ EARLIER APPLICATION NUMBER: 08/427,023
/ EARLIER FILING DATE: 1995-04-24
/ NUMBER OF SEQ ID NOS: 66
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 48
/ TYPE: PRT
/ LENGTH: 182
/ ORGANISM: Borrelia afzelii
US-09-117-257-48

Query Match      44.4%; Score 36; DB 3; Length 182;
Best Local Similarity 47.1%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPVXNA 19
Db 139 EKTPTTAAEGITIAKA 155

RESULT 33
US-09-489-352-48
/ Sequence 48, Application US/09489352
/ Patent No.: 6312907
/ GENERAL INFORMATION:
/ APPLICANT: Hook, Magnus
/ APPLICANT: Guo, Betty
/ APPLICANT: Hanson, Mark
/ TITLE OF INVENTION: DcpA AND DcpB COMPOSITIONS AND METHODS OF USE
/ FILE REFERENCE: 4210.000500
/ CURRENT APPLICATION NUMBER: US/09/489,352
/ CURRENT FILING DATE: 2000-01-21
/ EARLIER APPLICATION NUMBER: PCT/US96/17081
/ EARLIER FILING DATE: 1996-10-22
/ EARLIER APPLICATION NUMBER: 08/589,711
/ EARLIER FILING DATE: 1996-01-22
/ EARLIER APPLICATION NUMBER: 08/427,023
/ EARLIER FILING DATE: 1995-04-24
/ NUMBER OF SEQ ID NOS: 66
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 48
/ TYPE: PRT
/ LENGTH: 182
/ ORGANISM: Borrelia afzelii
US-09-489-352-48

Query Match      44.4%; Score 36; DB 4; Length 182;
Best Local Similarity 47.1%; Pred. No. 99;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPVXNA 19
Db 139 EKTPTTAAEGITIAKA 155

RESULT 34
US-09-134-001C-3558
/ Sequence 3558, Application US/09134001C
/ Patent No.: 6380370
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
/ FILE REFERENCE: GTC-007
/ CURRENT APPLICATION NUMBER: US/09/134,001C
/ CURRENT FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/064,964
/ PRIOR FILING DATE: 1997-11-08
/ PRIOR APPLICATION NUMBER: US 60/055,779
/ PRIOR FILING DATE: 1997-08-14

/ NUMBER OF SEQ ID NOS: 5674
/ SEQ ID NO 3558
/ LENGTH: 236
/ TYPE: PRT
/ ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3558

Query Match      44.4%; Score 36; DB 4; Length 236;
Best Local Similarity 46.7%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 KTPLTTAAAPVXN 18
Db 65 KAPITLVALPVVEN 79

RESULT 35
US-09-446-301A-6
/ Sequence 6, Application US/09446301A
/ Patent No.: 6508893
/ GENERAL INFORMATION:
/ APPLICANT: EL SOLE, NEVINE
/ APPLICANT: ALIGNET, JEANINE
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
/ TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
/ TITLE OF INVENTION: COMPOUNDS
/ FILE REFERENCE: 03715-0059
/ CURRENT APPLICATION NUMBER: US/09/446,301A
/ CURRENT FILING DATE: 1999-12-20
/ NUMBER OF SEQ ID NOS: 51
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 294
/ TYPE: PRT
/ ORGANISM: Staphylococcus sp.
US-09-446-301A-6

Query Match      44.4%; Score 36; DB 4; Length 294;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPV 15
Db 173 EYEPLEPTNAAPV 185

RESULT 36
US-09-099-932-6
/ Sequence 6, Application US/09099932
/ Patent No.: 6570001
/ GENERAL INFORMATION:
/ APPLICANT: El Solh, Nevine
/ APPLICANT: ALIGNET, JEANINE
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
/ TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
/ TITLE OF INVENTION: COMPOUNDS
/ FILE REFERENCE: 03495.0173-00000
/ CURRENT APPLICATION NUMBER: US/09/099,932
/ CURRENT FILING DATE: 1998-06-19
/ EARLIER APPLICATION NUMBER: 60/050,380
/ EARLIER FILING DATE: 1997-06-20
/ NUMBER OF SEQ ID NOS: 50
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 295
/ TYPE: PRT
/ ORGANISM: Staphylococcus
US-09-099-932-6

Query Match      44.4%; Score 36; DB 4; Length 295;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```


MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/821,717B
FILING DATE: 15-JAN-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Timain, Susan J.
REGISTRATION NUMBER: 34,103
REFERENCE/DOCKET NUMBER: 20884/21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1636
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PUBLICATION INFORMATION:
AUTHORS: Lopez, Jose A.
AUTHORS: Chung, Dominic W.
AUTHORS: Fujikawa, Kazuo
AUTHORS: Hagen, Frederick S.
AUTHORS: Papayannopoulou, Thalia
AUTHORS: Roth, Gerald J.
TITLE: Cloning of the alpha chain of human
TITLE: platelet glycoprotein Ib: A transmembrane protein with homology
TITLE: to leucine-rich alpha-2-glycoprotein
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 84
PAGES: 5615-5619
DATE: AUG-1987
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 610
PUBLICATION INFORMATION:
AUTHORS: Zimmerman, Theodore S.
AUTHORS: Ruggeri, Zaverio M.
AUTHORS: Houghten, Richard A.
AUTHORS: Vincete, Vincete
AUTHORS: Mohri, Hiroshi
TITLE: Proteolytic fragments and synthetic
TITLE: peptides that block the binding of von Willebrand factor to the
TITLE: platelet membrane glycoprotein Ib
DOCUMENT NUMBER: EP 0 317 278 A2
FILING DATE: 16-NOV-1988
PUBLICATION DATE: 24-MAY-1989
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 0 TO 293
US-07-821-717B-6

Query Match 44.4%; Score 36; DB 1; Length 610;
Best Local Similarity 42.9%; Pred. NO. 3.9e+02;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPV 16
DB 409 EPTPIATISPTI 422

Search completed: October 4, 2004, 15:18:35
Job time: 31 secs

B/ank

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 15:16:39 ; Search time 108 Seconds
(Without alignments)
56.613 Million cell updates/sec

Title: US-09-359-426C-2

Perfect score: 81

Sequence: 1 XEEXTPLTAAAPVWVNA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53	65.4	495	12	US-10-282-122A-68674
2	48	59.3	480	12	US-10-282-122A-77773
3	45	55.6	478	15	US-10-369-493-13806
4	44	54.3	104	15	US-10-437-963-19808
5	44	54.3	306	14	US-10-238-075-1139
6	44	54.3	584	12	US-10-282-122A-47689
7	43	53.1	373	16	US-10-437-963-141172
8	43	53.1	483	14	US-10-156-761-10589
9	42	51.9	322	12	US-10-425-114-63257
10	41	50.6	132	13	US-10-078-929-22
11	41	50.6	132	13	US-10-078-929-198
12	41	50.6	132	16	US-10-437-963-14967
13	41	50.6	330	16	US-10-767-701-42700
14	41	50.6	517	12	US-10-282-122A-51339
15	40	49.4	172	16	US-10-767-701-57843

ALIGNMENTS

RESULT 1

US-10-282-122A-68674
; Sequence 68674, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carl, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636

16	40	49.4	346	14	US-10-156-761-13589	Sequence 13589, A
17	40	49.4	443	9	US-09-815-242-11701	Sequence 11701, A
18	40	49.4	498	9	US-09-886-156-59	Sequence 59, Appl
19	40	49.4	498	9	US-09-886-150-59	Sequence 59, Appl
20	40	49.4	498	10	US-09-886-149-59	Sequence 59, Appl
21	40	49.4	498	10	US-09-886-153-59	Sequence 59, Appl
22	40	49.4	498	14	US-10-326-090-59	Sequence 59, Appl
23	40	49.4	500	14	US-10-166-087-4	Sequence 138004, A
24	40	49.4	1133	16	US-10-437-963-138004	Sequence 138004, A
25	39	48.1	61	16	US-10-424-599-214148	Sequence 214148, A
26	39	48.1	106	12	US-09-952-432A-2	Sequence 2, Appl
27	39	48.1	155	9	US-09-952-432A-16	Sequence 16, Appl
28	39	48.1	233	9	US-09-952-432A-2	Sequence 5663, Ap
29	39	48.1	283	12	US-10-424-599-167461	Sequence 167461, A
30	39	48.1	289	16	US-10-437-963-203676	Sequence 203676, A
31	39	48.1	322	16	US-10-437-963-112045	Sequence 112045, A
32	39	48.1	387	9	US-09-952-432A-19	Sequence 19, Appl
33	39	48.1	395	9	US-09-952-432A-21	Sequence 21, Appl
34	39	48.1	477	15	US-10-369-493-8852	Sequence 8852, Ap
35	39	48.1	484	12	US-10-282-122A-68665	Sequence 68665, A
36	39	48.1	484	12	US-10-406-686A-29	Sequence 29, Appl
37	39	48.1	507	15	US-10-369-493-16863	Sequence 16863, A
38	39	48.1	507	16	US-10-389-566-1128	Sequence 1128, Ap
39	39	48.1	509	14	US-10-156-761-10971	Sequence 10971, A
40	39	48.1	1353	15	US-10-369-493-5890	Sequence 5890, Ap
41	39	48.1	2080	15	US-10-353-690-38	Sequence 36, Appl
42	38	46.9	75	12	US-10-424-599-157296	Sequence 157296, A
43	38	46.9	101	16	US-10-437-963-158367	Sequence 158367, A
44	38	46.9	116	12	US-10-425-114-68517	Sequence 68517, A
45	38	46.9	116	12	US-10-425-114-68517	Sequence 68517, A

```

; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68674
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-282-122A-68574

```

```
Query Match      65.4%; Score 53; DB 12; Length 495;
Best Local Similarity 75.0%; Pred. No. 1.2;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

Qy 3 EKTPLTTAAXAPVXN 18
 || | | | | | |
Db 13 EKKKLTTAAGAPVDN 28

RESULT 2
US-10-282-122A-77773
; sequence 77773, Application US/10282122A
; Publication No. US20040029129A1

Query Match 59.3%; Score 48; DB 12; Length 480;
Best Local Similarity 68.8%; Pred. No. 8.7;

```

Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      3  EKTPLTTAAXAPVXXN 18
      :  ||||| |||||
Db       4  KKKGLTTAAGAPVDN 19

RESULT 3
US-10-369-493-13806
; Sequence 13806, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13806
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Pseudomonas fluorescens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(478)
; OTHER INFORMATION: unsure at all xaa locations
US-10-369-493-13806

```

Query Match 55.6%; Score 45; DB 15; Length 478;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LTAAAXAPVXN 18
Db 1 LTAAAGAPVDN 12

```

RESULT 4
US-10-437-963-199308
; Sequence 199308, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbasz, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid MM
; FILE REFERENCE: Plants and Uses Th
; CURRENT APPLICATION NUMBER: US/10/437,9
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 199308
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURES:
; US-10-437-963-199308
; OTHER INFORMATION: Clone ID: PAT_MRT145

```

Query Match 54.3%; Score 44; DB 16; Length 104;
Best Local Similarity 69.2%; Pred. No. 7.6;

```
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 EKPTLTAAAXAP 14
DB 79 EKPTLVTPOGAP 91
RESULT 5
US-10-238-075-1139
; Sequence 1139, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolated
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1139
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-238-075-1139
Query Match 54.3%; Score 44; DB 14; Length 306;
Best Local Similarity 47.1%; Pred. No. 26;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 3 EKPTLTAAAXAPVXNA 19
DB 84 EKVPCTSSGVAFVNA 100
RESULT 6
US-10-282-122A-47689
; Sequence 47689, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
```

```
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47689
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-47689
Query Match 54.3%; Score 44; DB 12; Length 584;
Best Local Similarity 56.2%; Pred. No. 53;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 4 KPTLTAAAXAPVXNA 19
DB 101 RTPSLSTAAQVVYNA 116
RESULT 7
US-10-437-963-141172
; Sequence 141172, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 141172
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_42301C.1.pep
US-10-437-963-141172
Query Match 53.1%; Score 43; DB 16; Length 373;
Best Local Similarity 61.5%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 2 EKPTLTAAAXAP 14
DB 222 DEETPTTAAABP 234
RESULT 8
US-10-156-761-10589
; Sequence 10589, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
```

; CURRENT APPLICATION NUMBER: US/10/156.761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10589
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10589

Query Match 53.1%; Score 43; DB 14; Length 483;
Best Local Similarity 69.2%; Pred. No. 63;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 PLTTAAAXAPVXN 18
| | | | | | | | | | | | | | | | | | | | | |
Db 5 PLTTERGAPVADN 17

RESULT 9
US-10-425-114-63257
; Sequence 63257, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63257
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3180-051-E8_FLI.pep
US-10-425-114-63257

Query Match 51.9%; Score 42; DB 12; Length 322;
Best Local Similarity 50.0%; Pred. No. 60;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 BEKPLTTAAAXAPVXNA 19
| | | | | | | | | | | | | | | | | | | | | |
Db 44 BEPTTAATAASPAPDRA 61

RESULT 10
US-10-078-929-22
; Sequence 22, Application US/10078929
; Publication No. US20020152497A1
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Antoni
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Sakai, Hajime
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Odell, Joan T.
; APPLICANT: Meyers, Blake
; APPLICANT: Thorpe, Catherine
; TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
; FILE REFERENCE: BB1357 US NA

; CURRENT APPLICATION NUMBER: US/10/078.929
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/566,394
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/133038
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133042
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133427
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133437
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133428
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133438
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133436
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/137667
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-078-929-22

Query Match 50.6%; Score 41; DB 13; Length 132;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 BEKPLTTAAAXAPVXNA 19
| | | | | | | | | | | | | | | | | | | | | |
Db 97 EAAPTTTAAEAPATAA 114

RESULT 11
US-10-078-929-198
; Sequence 198, Application US/10078929
; Publication No. US20020152497A1
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Antoni
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Sakai, Hajime
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Odell, Joan T.
; APPLICANT: Meyers, Blake
; APPLICANT: Thorpe, Catherine
; TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
; FILE REFERENCE: BB1357 US NA
; CURRENT APPLICATION NUMBER: US/10/078.929
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/566,394
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/133038
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133042
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133427
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133437
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133428
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133438
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133436
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/137667


```
; Sequence 57843, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 57843
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(172)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: 30972933.pap
; US-10-767-701-57843

Query Match          49.4%; Score 40; DB 16; Length 172;
Best Local Similarity 88.9%; Pred. No. 65;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 PLTTAAXAP 14
   |||||
Db 51 PLTTAAAP 59

RESULT 16
; US-10-156-761-13589
; Sequence 13589, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13589
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
; US-10-156-761-13589

Query Match          49.4%; Score 40; DB 14; Length 346;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 TPLTTAAXAPVXNA 19
   |||||
Db 16 TVLTTAAAVVLSA 30

RESULT 17
; US-09-815-242-11701
; Sequence 11701, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11701
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-815-242-11701

Query Match          49.4%; Score 40; DB 9; Length 445;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 EKTPLTTAAXAPV 15
   :|||
Db 382 EKTPLTAADMAAI 395

RESULT 18
; US-09-886-156-59
; Sequence 59, Application US/09886156
; Patent No. US20020155428A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886,156
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US/09/109,916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
```

```

; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-156-59

Query Match          49.4%; Score 40; DB 9; Length 498;
Best Local Similarity 44.4%; Pred. No. 2.2e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 EEXTPLTTAAXAPVVXNA 19
   |||:|:|:|:|:|:|
Db 119 EETSPTQTSQNYPIVTNA 136

RESULT 19
US-09-886-150-59
; Sequence 59, Application US/09886150
; Patent No. US20020172939A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886.150
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/109, 916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-150-59

Query Match          49.4%; Score 40; DB 9; Length 498;
Best Local Similarity 44.4%; Pred. No. 2.2e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 EEXTPLTTAAXAPVVXNA 19
   |||:|:|:|:|:|:|
Db 119 EETSPTQTSQNYPIVTNA 136

RESULT 20
US-09-886-149-59
; Sequence 59, Application US/09886149
; Publication No. US2003003442A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886.149
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/109, 916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-150-59

Query Match          49.4%; Score 40; DB 9; Length 498;
Best Local Similarity 44.4%; Pred. No. 2.2e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 EEXTPLTTAAXAPVVXNA 19
   |||:|:~:~:~:~:~:~:~
Db 119 EETSPTQTSQNYPIVTNA 136

RESULT 21
US-09-886-159-59
; Sequence 59, Application US/09886159
; Publication No. US20030003443A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886.159
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US/09/109.916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7
; PRIOR FILING DATE: 1992-10-22
; PRIOR APPLICATION NUMBER: DE P 42 44 541.8
; PRIOR FILING DATE: 1992-12-30
; PRIOR APPLICATION NUMBER: DE P 43 18 186.4
; PRIOR FILING DATE: 1993-06-01
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 59
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-09-886-159-59

Query Match          49.4%; Score 40; DB 10; Length 498;
Best Local Similarity 44.4%; Pred. No. 2.2e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 EEXTPLTTAAXAPVVXNA 19
   |||:|:~:~:~:~:~:~:~
Db 119 EETSPTQTSQNYPIVTNA 136

RESULT 22
US-10-326-090-59
; Sequence 59, Application US/10326090
; Publication No. US20030166915A1
; GENERAL INFORMATION:
; APPLICANT: Guertler, Lutz G.
; APPLICANT: Eberle, Josef
; APPLICANT: Brunn, Albrecht V.
; APPLICANT: Knapp, Stefan
; APPLICANT: Hauser, Hans-Peter
; TITLE OF INVENTION: RETROVIRUS FROM HIV GROUP AND ITS USE
; FILE REFERENCE: 05495.0001-04
; CURRENT APPLICATION NUMBER: US/09/886.149
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/109, 916
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: DE P 42 33 646.5
; PRIOR FILING DATE: 1992-10-06
; PRIOR APPLICATION NUMBER: DE P 42 35 718.7

```


; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 214148
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(106)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_35402C.1.pep
US-10-424-599-214148

Query Match 48.1%; Score 39; DB 12; Length 106;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 EKTPLTTAAAPV 14
||| ||| ||| |||
Db 86 EKTPLNTVDEAP 97

RESULT 27
US-09-952-432A-2
; Sequence 2, Application US/09952432A
; Patent No. US20020150588A1
; GENERAL INFORMATION:
; APPLICANT: Allison, James P.
; APPLICANT: Fasso, Marcella
; APPLICANT: Shastrri, Nilabh
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: SPAS-1 CANCER ANTIGEN
; FILE REFERENCE: 018941-001100S
; CURRENT APPLICATION NUMBER: US/09/952,432A
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 60/234,472
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SPAS-1 cDNA
US-09-952-432A-2

Query Match 48.1%; Score 39; DB 9; Length 155;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 TPLTTAAAPV 16
: ||| ||| ||| |||
Db 60 SPTTTAATWV 71

RESULT 28
US-09-952-432A-16
; Sequence 16, Application US/09952432A
; Patent No. US20020150588A1
; GENERAL INFORMATION:
; APPLICANT: Allison, James P.
; APPLICANT: Fasso, Marcella
; APPLICANT: Shastrri, Nilabh
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: SPAS-1 CANCER ANTIGEN
; FILE REFERENCE: 018941-001100S
; CURRENT APPLICATION NUMBER: US/09/952,432A
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 60/234,472
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Predicted AA
; OTHER INFORMATION: sequence of mouse SPAS-1
US-09-952-432A-16

Query Match 48.1%; Score 39; DB 9; Length 155;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 TPLTTAAAPV 16
: ||| ||| ||| |||
Db 60 SPTTTAATWV 71

RESULT 29
US-09-738-626-5663
; Sequence 5663, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 5663
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-5663

Query Match 48.1%; Score 39; DB 9; Length 233;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 TPLTTAAAPV 16
||| ||| ||| |||
Db 35 TPTTSASPAPV 46

RESULT 30
US-10-424-599-167461
; Sequence 167461, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B

```

; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 167461
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MET3847
US-10-424-599-167461

```

Query Match 48.1%; Score 39; DB 12; Length 265;
Best Local Similarity 72.7%; Pred. No. 1.6e+02;
Matches 8; Conservative 1; Mismatches 2; Indels

Qy	5	TPLTTAAXAPV	15
			:
D _b	201	TPLTRAAHAPM	211

```

RESULT 31
US-10-437-963-203676
; Sequence 203676, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Bing
; TITLE OF INVENTION: Rice Nucleic Acid Molecules
; TITLE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 203676
; LENGTH: 269
; TYPE: PRI
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRI451
US-10-437-963-203676

```

```
Query Match      48.1%;   Score 39;   DB 16;   Length 269;
Best Local Similarity 50.0%;   Pred. No. 1.6e+02;
Matches      8;   Conservative      2;   Mismatches      6   Indels
```

Qy 4 KTPLTTAAXAPVXNA 19
|||:|:|
Dp 221 KTPATSVADAPEADA 236

RESULT 32
US-10-437-963-112045
/ Sequence 112045, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid M
/ TITLE OF INVENTION: Plants and Uses Th
/ FILE REFERENCE: 39-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,9

```

; CURRENT FILING DATE: 2003-05-14
;
; NUMBER OF SEQ ID NOS: 204966
;
; SEQ ID NO 112045
;
; LENGTH: 387
;
; TYPE: PRT
;
; ORGANISM: Oryza sativa
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15968C.1.pep
US-10-437-963-112045

```

Query Match 48.1%; Score 39; DB 16; Length 387;
Best Local Similarity 41.2%; Pred. NO. 2.4e+02;
Matches 7; Conservative 5; Mismatches 5; Indels

```
Qy      2 EEKPLTTAAAXAPVVXN 18
        ::||: ||||| :
Db     370 KKETPVIVAAAPAVAS 380
```

RESULT 33
US-09-952-432A-19
; Sequence 19, Application US/09952432A
; Patent No. US2002015058A1
; GENERAL INFORMATION:
; APPLICANT: Allison, James P.
; APPLICANT: Fasso, Marcella
; APPLICANT: Shastri, Nilabh
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: SPAS-1 CANCER ANTIGEN
; FILE REFERENCE: 018941-001110US
; CURRENT APPLICATION NUMBER: US/09/952,432A
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 60/234,472
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Tumor SPAS-1
US-09-952-432A-19

Query Match 48.1%; Score 39; DB 9; Length 395;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 3; Indels

QY 5 TPLTTAAXAPVV 16
:||||
Db 300 SPTTTAATMPVV 311

```

RESULT 34
US-09-952-432A-21
; Sequence 21, Application US/09952432A
; Patent No. US2002015058A1
; GENERAL INFORMATION:
; APPLICANT: Allison, James P.
; APPLICANT: Passo, Marcela
; APPLICANT: Shastri, Nilabh
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: SPAS-1 CANCER ANTIGEN
; FILE REFERENCE: 018941-001110US
; CURRENT APPLICATION NUMBER: US/09/952,432A
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: 60/234,472
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 395
; TYPE: PRT

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020150588A1mal SPAS-1
US-09-952-432A-21

Query Match      48.1%; Score 39; DB 9; Length 395;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TPLTTAAXAPVV 16
   : |||||
Db 300 SPTTAAATMPVV 311

RESULT 35
US-10-369-493-8852
; Sequence 3852, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8852
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Ralstonia metallidurans
US-10-369-493-8852

Query Match      48.1%; Score 39; DB 15; Length 477;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 LTTAAXAPVVXN 18
   |||||
Db 1 LTTAFGAPVDN 12

RESULT 36
US-10-282-122A-66865
; Sequence 66865, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-03-23
; PRIOR APPLICATION NUMBER: 60/207,727

```

```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66865
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-282-122A-66865

Query Match      48.1%; Score 39; DB 12; Length 484;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 LTTAAXAPVVXN 18
   |||||
Db 13 LTNAGAPIVEN 24

RESULT 37
US-10-406-686A-29
; Sequence 29, Application US/10406686A
; Publication No. US2004003386A1
; GENERAL INFORMATION:
; APPLICANT: CROCKE, HELEN RACHEL
; APPLICANT: SHEA, JACQUELINE ELIZABETH
; APPLICANT: FELDMAN, ROBERT GRAHAM
; APPLICANT: COUTEBROZE, SYLVAIN GABRIEL
; APPLICANT: LEGROS, FRANCOIS-XAVIER
; TITLE OF INVENTION: ATTENUATED GRAM NEGATIVE BACTERIA
; FILE REFERENCE: 454313-3171.1
; CURRENT APPLICATION NUMBER: US/10/406,686A
; CURRENT FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/370,282
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-10-406-686A-29

Query Match      48.1%; Score 39; DB 12; Length 484;
Best Local Similarity 66.7%; Pred. No. 3.1e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 LTTAAXAPVVXN 18
   |||||
Db 13 LTNAGAPIVEN 24

RESULT 38
US-10-369-493-16863
; Sequence 16863, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.

```

; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 16863
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-10-369-493-16863

Query Match 48.1%; Score 39; DB 15; Length 507;
Best Local Similarity 47.1%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPVXNA 19
||||| :
Db 185 EKTPLTALATQAILERA 201

RESULT 39
US-10-389-566-1128
; Sequence 1128, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1128
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Caulobacter crescentus
US-10-389-566-1128

Query Match 48.1%; Score 39; DB 16; Length 507;
Best Local Similarity 47.1%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPVXNA 19
||||| :
Db 185 EKTPLTALATQAILERA 201

RESULT 40
US-10-156-761-10971
; Sequence 10971, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10971
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10971

Query Match 48.1%; Score 39; DB 14; Length 509;
Best Local Similarity 52.9%; Pred. No. 3.3e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPVXNA 19
||||| :
Db 189 EITPLTTAALVDLIVEA 205

Search completed: October 4, 2004, 15:27:56
Job time : 110 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 14:57:29 ; Search time 71 Seconds
(without alignments)
25.741 Million cell updates/sec

Title: US-09-359-426C-2

Perfect score: 81

Sequence: 1 XEKTPLTTAAXAPVXNA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	77.8	482	2 B83113	catalase PA4236 [i
2	53	65.4	484	2 A58663	catalase (EC 1.11.
3	48	59.3	480	2 A30148	catalase (EC 1.11.
4	46	56.8	262	2 T33408	hypothetical prote
5	46	56.8	262	2 E88400	protein H34124.2 [
6	45	55.6	128	2 JCL273	ribosomal protein
7	44	54.3	306	2 A80963	probable carbonyl
8	44	54.3	427	2 F85436	hypothetical prote
9	43	53.1	483	2 S37055	catalase (EC 1.11.
10	42	51.9	105	2 A44639	catalase (EC 1.11.
11	42	51.9	321	2 T08462	hypothetical prote
12	42	51.9	436	2 T46107	hypothetical prote
13	41	50.6	132	2 T50779	copper chaperone h
14	41	50.6	482	2 S60757	catalase (EC 1.11.
15	40	49.4	171	2 T31478	hypothetical prote
16	40	49.4	252	2 H72469	hypothetical prote
17	40	49.4	258	2 A01328	hypothetical prote
18	40	49.4	394	2 E87611	hypothetical prote
19	40	49.4	788	2 S05661	muscarinic acetyl
20	40	49.4	1360	2 T31674	hypothetical prote
21	39.5	48.8	405	2 J02147	OHPI protein - mai
22	39	48.1	486	2 S10772	2-hydroxymuconic s
23	39	48.1	507	2 B87400	aldehyde dehydroge
24	39	48.1	527	2 S46088	hypothetical prote
25	39	48.1	1353	2 T56301	hypothetical prote
26	38.5	47.5	1658	2 D75489	hypothetical prote
27	38	46.9	143	2 S68226	growth-blocking pe
28	38	46.9	307	2 T40089	cytochrome c1, hem
29	38	46.9	307	2 A35125	phospholipase D (E

30	38	46.9	348	2 S40750	hypothetical prote
31	38	46.9	474	2 E87650	peptidase, M20/M25
32	38	46.9	532	2 S40983	hypothetical prote
33	38	46.9	573	2 T81313	peptidase (M3 faml
34	38	46.9	684	2 T25603	hypothetical prote
35	38	46.9	695	2 D71283	probable translati
36	38	46.9	747	2 F88561	protein F88A4.11 [
37	38	46.9	1679	2 S48385	hypothetical prote
38	37.5	46.3	200	2 G86194	hypothetical prote
39	37	45.7	71	2 F72332	hypothetical prote
40	37	45.7	116	2 C83492	hypothetical prote
41	37	45.7	328	2 T01225	hypothetical prote
42	37	45.7	347	2 T35518	probable secreted
43	37	45.7	463	2 S72992	probable phosphory
44	37	45.7	488	2 T42038	catalase (EC 1.11.
45	37	45.7	500	2 S72505	lycopenase beta-cycl

ALIGNMENTS

RESULT 1

B83113 catalase PA4236 [imported] - Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C/Accession: B83113

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A/Reference number: A82950; MUID:20437337; PMID:10984043

A/Accession: B83113

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-482 <STO>

A/Cross-references: GB:AE004841; GB:AE004091; NID:G9950451; PIDN:AAG07624.1; GSPDB:GN001

A/Experimental source: strain PA01

C/Genetics:

A/Gene: katA; PA4236

C/Superfamily: catalase

Query Match 77.8%; Score 63; DB 2; Length 482;

Best Local Similarity 82.4%; Pred. No. 0.0041; 3; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 EKTPLTTAAXAPVXN 18

|||||

Db 2 EKTPLTTAAXAPVXN 18

RESULT 2

A58663 catalase (EC 1.11.1.6) [validated] - Proteus mirabilis

C/Species: Proteus mirabilis

A/Variety: strain P, peroxide resistant

C/Date: 19-Nov-1997 #sequence_revision 21-Nov-1997 #text_change 15-Sep-2000

C/Accession: A58663; B58663

R/Buzy, A.; Bracchi, V.; Sterjades, R.; Chroboczek, J.; Thibault, P.; Gagnon, J.; Jouve,

J. Protein Chem. 14, 59-72, 1995

A/Title: Complete amino acid sequence of Proteus mirabilis PR catalase. Occurrence of a n

A/Reference number: A58663; MUID:95305957; PMID:7786407

A/Accession: A58663

A/Molecule type: protein

A/Residues: 1-484 <BUZ1>

A/Accession: B58663

A/Molecule type: DNA

A/Residues: 1-305, 'AE' <BUZ2>

R/Guuet, P.; Jouve, H.M.; Dideberg, O.

submitted to the Brookhaven Protein Data Bank, June 1996

A/Reference number: A57899; PDB:2CAE

A/Contents: annotation; X-ray crystallography, 2.2 angstroms, residues 3-475

R.Gouet, P.; Jouve, H.M.; Hajdu, J.
 Submitted to the Brookhaven Protein Data Bank, June 1996
 A:Reference number: A67900; PDB:2CAF
 A:Contents: annotation; X-ray crystallography, 2.7 angstroms, residues 3-475
 A:Note: compound I
 R.Gouet, P.; Jouve, H.M.; Hajdu, J.
 Submitted to the Brookhaven Protein Data Bank, June 1996
 A:Reference number: A67901; PDB:2CAG
 A:Contents: annotation; X-ray crystallography, 2.7 angstroms, residues 3-475
 A:Note: compound II, dithiothreitol reduced compound I
 R.Gouet, P.; Jouve, H.M.; Dideberg, O.
 Submitted to the Brookhaven Protein Data Bank, July 1996
 A:Reference number: A67902; PDB:2CAH
 A:Contents: annotation; X-ray crystallography, 2.7 angstroms, residues 3-475
 A:Note: native Pe(III) with NADPH
 R.Gouet, P.; Jouve, H.M.; Dideberg, O.
 J. Mol. Biol. 249, 933-954, 1995
 A:Title: Crystal structure of proteus mirabilis PR catalase with and without bound NADPH
 A:Reference number: A58664; MUID:95311317; PMID:7791219
 A:Contents: annotation; X-ray crystallography, 2.2 angstroms
 A:Complex: homotetramer
 C:Function:
 A:Description: catalyzes the conversion of two of molecules of hydrogen peroxide to two
 A:Note: this enzyme has a tightly bound NADPH cofactor
 C:Superfamily: catalase
 C:Keywords: chromoprotein; heme; homotetramer; iron; metalloprotein; NADP; oxidoreductase
 F:53/Modified site: methionine sulfone (Met) #status experimental
 F:54.93.127/Active site: His, Ser, Asn #status predicted
 F:7337/Binding site: heme iron (Tyr) (axial ligand) #status experimental

Query Match 65.4%; Score 53; DB 2; Length 484;
 Best Local Similarity 75.0%; Pred. No. 0.22;
 Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 EKTPLTTAAXAPVYXN 18
 |||||:|||||
 DB 2 EKKLLTTAAGAPVVDN 17
 |||||:|||||

RESULT 3
 AB0148
 catalase (EC 1.11.1.6) [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AB0148
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11566360
 A:Accession: AB0148
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-480 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CA90045.1; PID:gl5979266; GSPDB:GN00175
 C:Genetics:
 A:Gene: katA
 C:Superfamily: catalase
 C:Keywords: oxidoreductase

Query Match 59.3%; Score 48; DB 2; Length 480;
 Best Local Similarity 68.8%; Pred. No. 1.6;
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 EKTPLTTAAXAPVYXN 18
 :|||||:|||||
 DB 4 KKKGLTTAAGAPVVDN 19
 :|||||:|||||

RESULT 4
 T33408
 hypothetical protein H34I24.2 - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T33408
 R:Latreille, P.; Wamsley, P.; O'Brien, D.
 submitted to the EMBL Data Library, July 1998
 A:Description: The sequence of C. elegans cosmid H34I24.
 A:Reference number: Z21340
 A:Accession: T33408
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-262 <LAT>
 A:Cross-references: EMBL:AF078784; PIDN:AAC26925.1; GSPDB:GN00021; CESP:H34I24.2
 A:Experimental source: strain Bristol N2; clone H34I24
 C:Genetics:
 A:Gene: CESP:H34I24.2
 A:Map position: 3
 A:Note: intron positions not resolved (incomplete sequence)

Query Match 56.8%; Score 46; DB 2; Length 262;
 Best Local Similarity 56.2%; Pred. No. 1.9;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 KTPLTTAAXAPVYXNA 19
 |||||:|||||
 DB 190 KTPLTSGSSARVINNA 205
 |||||:|||||

RESULT 5
 E88400
 protein H34I24.2 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: E88400
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: E88400
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-262 <STO>
 A:Cross-references: GB:chr_III; PIDN:AAC26925.1; PID:g3329614; GSPDB:GN00021; CESP:H34I24.2
 C:Genetics:
 A:Map position: 3

Query Match 56.8%; Score 46; DB 2; Length 262;
 Best Local Similarity 56.2%; Pred. No. 1.9;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 KTPLTTAAXAPVYXNA 19
 |||||:|||||
 DB 190 KTPLTSGSSARVINNA 205
 |||||:|||||

RESULT 6
 JC1273
 ribosomal protein L7/L12 - Streptomyces antibioticus
 C:Species: Streptomyces antibioticus
 C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Aug-1999
 C:Accession: JC1273
 R:Parra, F.; Blanco, G.; Alonso, J.M.M.; Balbin, M.; Mendez, C.; Salas, J.A.
 Gene 118, 127-129, 1992
 A:Title: Cloning and sequence of a gene encoding the L7/L12 ribosomal protein equivalent
 A:Reference number: JC1273; MUID:92380478; PMID:1511874
 A:Accession: JC1273
 A:Molecule type: DNA
 A:Residues: 1-128 <PAR>
 A:Cross-references: GB:M89911; MUID:gl53436; PIDN:AAA26611.1; PID:gl53438
 C:Superfamily: Escherichia coli ribosomal protein L12
 C:Keywords: protein biosynthesis; ribosome

Query Match 55.6%; Score 45; DB 2; Length 128;
 Best Local Similarity 66.7%; Pred. No. 1.3;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 BEKTPLTAAAPV 16
 ||||:|||||
 Db 28 BEKFDVTAAPV 42

RESULT 7
 AE0963
 Probable carbohydrate kinase STY3989 [imported] - Salmonella enterica subsp. enterica serovar Typhi
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AE0963
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AE0963
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-306 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD03201.1; PID:gl6504832; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY3989
 C:Superfamily: ribokinase

Query Match 54.3%; Score 44; DB 2; Length 306;
 Best Local Similarity 47.1%; Pred. No. 5;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 EKTPLTAAAPV 19
 ||||:|||||
 Db 84 EKVPCTSSGAPV 100

RESULT 8
 F85436
 hypothetical protein AT4G36970 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 C:Accession: F85436
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999
 A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A:Reference number: AB5001; MUID:20083488; PMID:10617198
 A:Accession: F85436
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-427 <STO>
 A:Cross-references: GB:NC_001268; NID:g7270646; PIDN:CAB80363.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4G36970
 A:Map position: 4

Query Match 54.3%; Score 44; DB 2; Length 427;
 Best Local Similarity 60.0%; Pred. No. 7.1;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 BEKTPLTAAAPV 16
 ||||:|||||
 Db 208 BEKAEVTTAQSPV 222

RESULT 9
 S37055
 catalase (EC 1.11.1.6) alpha-2 chain - Streptomyces violaceus

C:Species: Streptomyces violaceus
 C:Date: 10-Dec-1993 #sequence_revision 26-May-1995 #text_change 04-Mar-2000
 C:Accession: S37055
 R:Facey, S.; van Pee, K.H.; Vining, L.C.
 submitted to the EMBL Data Library, August 1993
 A:Reference number: S37055
 A:Accession: S37055
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-483 <PAC>
 A:Cross-references: EMBL:X74791; NID:g397888; PIDN:CAAS2796.1; PID:g581780
 A:Note: the source is designated as Streptomyces venezuelae
 C:Genetics:
 A:Start codon: GTG
 C:Superfamily: catalase
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
 F:54,93,127/Active site: His, Ser, Asn #status Predicted
 F:337/Binding site: heme iron (Tyr) (axial ligand) #status Predicted

Query Match 53.1%; Score 43; DB 2; Length 483;
 Best Local Similarity 69.2%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 PLTTAAXAPV 18
 ||||:|||||
 Db 5 PLTTAGAPV 17

RESULT 10
 A44639
 catalase (EC 1.11.1.6) - Streptomyces coelicolor (fragments)
 C:Species: Streptomyces coelicolor
 C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 31-Oct-1997
 C:Accession: A44639
 R:Walker, G.
 submitted to the Protein Sequence Database, September 1994
 A:Reference number: A44639
 A:Accession: A44639
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-105 <WAL>
 C:Superfamily: catalase
 C:Keywords: oxidoreductase

Query Match 51.9%; Score 42; DB 2; Length 105;
 Best Local Similarity 69.2%; Pred. No. 3.4;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 PLTTAAXAPV 18
 ||||:|||||
 Db 7 PLTTAGAPV 19

RESULT 11
 T08462
 hypothetical protein F2206.250 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 22-Oct-1999
 C:Accession: T08462
 R:Quetier, F.; Purnelle, B.; Boutry, M.; Goffeau, A.; Salanoubat, M.; Mewes, H.W.; Mayer, submitted to the Protein Sequence Database, May 1999
 A:Reference number: Z16420
 A:Accession: T08462
 A:Molecule type: DNA
 A:Residues: 1-321 <QUE>
 A:Cross-references: EMBL:AL050300; GSPDB:GN00061; ATSP:F2206.250
 A:Experimental source: cultivar Columbia; BAC clone F2206
 C:Genetics:
 A:Gene: ATSP:F2206.250
 A:Map position: 3

Query Match 51.9%; Score 42; DB 2; Length 321;
 Best Local Similarity 50.0%; Pred. No. 12;

QY 6 PLTTAAXAP 14
|||||
Db 64 PLTTAASAP 72

RESULT 17

AD1328
Hypothetical protein lmo2028 [imported] - Listeria monocytogenes (strain EGD-e)
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AD1328
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloesker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A/Authors: Kreitz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A./Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1328
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-258 <GLA>
A/Cross-references: GB:NC 003210; PIDN:CAD00106.1; PID:g16411498; GSPDB:GN00177
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo2028

Query Match 49.4%; Score 40; DB 2; Length 258;
Best Local Similarity 57.1%; Pred. No. 20;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 EKTPLTTAAXAPVV 16
|||||
Db 158 EETELTNAVYAPVI 171

RESULT 18

EB97611
Hypothetical protein CC2927 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C/Accession: EB97611
R./Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
n. J.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: AB87249; MUID:21173698; PMID:11259647
A/Accession: EB97611
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-394 <STG>
A/Cross-references: GB:AB005673; NID:gl3424551; PIDN:AAK24889.1; GSPDB:GN00148
C/Genetics:
A/Gene: CC2927

Query Match 49.4%; Score 40; DB 2; Length 394;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 KTLPLTTAAXAPVVXNA 19
|||||
Db 315 KTLPLTTTLTPALIDA 330

RESULT 19

S05661
muscarinic acetylcholine receptor - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 15-Oct-1999
C/Accession: S05661; A36191

R./Onai, T.; FitzGerald, M.G.; Arakawa, S.; Gosayne, J.D.; Urquhart, D.A.; Hall, L.M.; Fra
FEBS Lett. 255, 219-225, 1989
A/Title: Cloning, sequence analysis and chromosome localization of a Drosophila muscarini
A/Reference number: S05661; MUID:90005981; PMID:2507354
A/Accession: S05661
A/Molecule type: mRNA
A/Residues: 1-788 <GNA>
A/Cross-references: EMBL:M23412; NID:gl57829; PID:gl57830
R/Shapiro, R.A.; Wakimoto, B.T.; Subers, E.M.; Nathanson, N.M.
Proc. Natl. Acad. Sci. U.S.A. 86, 9039-9043, 1989
A/Title: Characterization and functional expression in mammalian cells of genomic and cDN
A/Reference number: A36191; MUID:90046926; PMID:2510174
A/Accession: A36191
A/Molecule type: mRNA
A/Residues: 81-154, T, 156-198, N, 200-215, G, 217-227, 229-330, P, 332-426, YATPVTIETPLQ
A/Cross-references: GB:M27495
A/Experimental source: strain Oregon R
C/Genetics:
A/Gene: FlyBase:mAcR-60C
A/Cross-references: FlyBase:FBgn0000037
A/Map position: 2R 60C5-C8
A/Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmem
F/104-127/Domain: transmembrane #status predicted <TM1>
F/141-161/Domain: transmembrane #status predicted <TM2>
F/179-200/Domain: transmembrane #status predicted <TM3>
F/220-243/Domain: transmembrane #status predicted <TM4>
F/268-289/Domain: transmembrane #status predicted <TM5>
F/702-722/Domain: transmembrane #status predicted <TM6>
F/739-758/Domain: transmembrane #status predicted <TM7>

Query Match 49.4%; Score 40; DB 2; Length 788;
Best Local Similarity 52.9%; Pred. No. 70;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 EKTPLTTAAXAPVVXNA 19
|||||
Db 602 EDGFTTTAAAPLAA 618

RESULT 20

T31574
Hypothetical protein 1 - fruit fly (Drosophila simulans) retrotransposon ninja (fragment)
C/Species: Drosophila simulans
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
C/Accession: T31574
R./Ogura, K.; Takechi, S.; Nakayama, T.; Yamamoto, M.
Genes Genet. Syst. 71, 1-8, 1996
A/Title: Molecular structure of the transposable element ninja in Drosophila simulans.
A/Reference number: Z21057; MUID:96252483; PMID:8925473
A/Accession: T31574
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-1360 <OGU>
A/Cross-references: EMBL:D83207; NID:dl068446; PID:dl020552; PIDN:BAAL9771.1
C/Genetics:
A/Cross-references: FlyBase:FBgn0015168
A/Mobile element: retrotransposon ninja

Query Match 49.4%; Score 40; DB 2; Length 1360;
Best Local Similarity 47.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 EKTPLTTAAXAPVVXNA 19
|||||
Db 101 EATPITTASQFPASSGA 117

RESULT 21

JQ2147
OHPI protein - maize
C/Species: Zea mays (maize)
C/Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 26-May-2000
C/Accession: JQ2147

submitted to the EMBL Data Library, February 1993

A:Reference number: S40750
A:Accession: S40750
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-348 <SMI>
A:Cross-references: EMBL:Z22173
C:Genetics:
A:Introns: 170/1; 215/3; 297/3; 343/1

Query Match 46.9%; Score 38; DB 2; Length 348;
Best Local Similarity 66.7%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TPLTTAAAPV 16
:|||||
DB 229 SPCTTAASAPSV 240

RESULT 31

EB7650
peptidase, M20/M25/M40 family [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: E87650
R:Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; R. J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4135-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: E87650
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-474 <STO>
A:Cross-references: GB:AE005673; NID:gl3424925; PIDN:AAK35201.1; GSPDB:GNC0148
C:Genetics:
A:Gene: CC3239

Query Match 46.9%; Score 38; DB 2; Length 474;
Best Local Similarity 57.1%; Pred. No. 88;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 PLTTAAAPVXNA 19
:|||||
DB 386 PLTPAIWAPIQKNA 399

RESULT 32

S40983
hypothetical protein F58A4.11 - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997
C:Accession: S40983
R:Berk, M.

submitted to the EMBL Data Library, February 1992

A:Reference number: S40973
A:Accession: S40983
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-532 <BER>
A:Cross-references: EMBL:Z22179; NID:g297956; PID:g297966
C:Genetics:
A:Introns: 82/3; 128/1; 222/1; 280/3; 397/2; 428/3; 491/2

Query Match 46.9%; Score 38; DB 2; Length 532;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TPLTTAAAPV 16
:|||||
DB 14 SPCTTAASAPSV 25

RESULT 33

F81313
peptidase (M3 family) Cj1099 [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: F81313
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A:Reference number: AB1250; MUID:20150912; PMID:10688204
A:Accession: F81313
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-573 <PAR>
A:Cross-references: GB:AL111168; NID:56968444; PIDN:CAB73354.1; PID:9696853;
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1099
C:Superfamily: oligoendopeptidase F

Query Match 46.9%; Score 38; DB 2; Length 573;
Best Local Similarity 61.5%; Pred. No. 1.1e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 EKTPLTTAAAPV 15
:|||||
DB 385 QNTPLTTAETASV 397

RESULT 34

T25603
hypothetical protein C32E8.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25603
R:Gattung, S.

submitted to the EMBL Data Library, February 1997

A:Description: The sequence of C. elegans cosmid C32E8.
A:Reference number: Z20056
A:Accession: T25603
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-884 <GAT>
A:Cross-references: EMBL:U88308; PIDN:AA842327.1; GSPDB:GN00019; CESP:C32E8.10
A:Experimental source: strain Bristol N2; clone C32E8
C:Genetics:
A:Gene: CESP:C32E8.10
A:Map position: 1
A:Introns: 24/2; 59/1; 163/3; 255/3; 282/3; 447/1; 487/1; 541/1; 587/2; 636/1

Query Match 46.9%; Score 38; DB 2; Length 684;
Best Local Similarity 44.4%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 2 EKTPLTTAAAPVXNA 19
:|||||
DB 487 EPENFTTAPAPQHNA 504

RESULT 35

D71283
probable translation elongation factor G (fusa-2) - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 02-Feb-2001
C:Accession: D71283
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinr
rson, J.; Kralak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDor
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876

A:Accession: D71283
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-695 <COL>
A:Cross-references: GB:AE001248; GB:AE000520; NID:g33233074; PIDN:AAC65735.1; PID:g33233074
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: IP0767
C:Superfamily: translation elongation factor G; translation elongation factor Tu homolog
C:Keywords: GTP binding; nucleotide binding; P-loop
F:3-139/Domain: translation elongation factor Tu homolog <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:136-139/Region: GTP-binding NKXD motif

Query Match 46.9%; Score 38; DB 2; Length 695;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 EEKPLTTAAAXAPVV 16
|||:|||||
Db 295 EEPVLTSTADAPVV 309
|||:|||||

RESULT 36
F88561
protein F58A4.11[imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: F88561
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: F88561
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-747 <STO>
A:Cross-references: GB:chr_III; PIDN:CAA80127.1; PID:g3874287; GSPDB:GN00021; CESP:F58A4
A:Gene: F58A4.11
A:Map position: 3

Query Match 46.9%; Score 38; DB 2; Length 747;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TPLTTAAAPVV 16
|||:|||||
Db 229 SPCTTAASAPV 240
|||:|||||

RESULT 37
S48385
hypothetical protein YIL149c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Apr-2002
C:Accession: S48385
R:Churcher, C.
submitted to the EMBL Data Library, September 1994
A:Reference number: S48310
A:Accession: S48385
A:Molecule type: DNA
A:Residues: 1-1679 <CHU>
A:Cross-references: GB:Z47047; EMBL:Z38059; NID:g603997; PID:g763197; GSPDB:GN00009; MIF
C:Genetics:
A:Gene: SGD.MLP2; MIPS.YIL149c
A:Cross-references: SGD:S000141
A:Map position: 9L

Query Match 46.9%; Score 38; DB 2; Length 1679;
Best Local Similarity 46.7%; Pred. No. 3.6e+02;

Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 EEKPLTTAAAXAPVV 16
|||:|||||
Db 346 EEBATMTTSAVPTV 360
|||:|||||

RESULT 38
G86194
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G86194
R:Rothberg, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.; ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G86194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-200 <STO>
A:Cross-references: GB:AE005172; NID:g8810460; PIDN:AAF80121.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 46.3%; Score 37.5; DB 2; Length 200;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 9; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

QY 2 EEKPLTTAAAXAPVVXNA 19
|||:|||||
Db 23 DELTFLAT--APVISSA 37
|||:|||||

RESULT 39
F72332
hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: F72332
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: F72332
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <ARN>
A:Cross-references: GB:AE001747; GB:AE000512; NID:g4981304; PIDN:AAD35863.1; PID:g498131
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0781

Query Match 45.7%; Score 37; DB 2; Length 71;
Best Local Similarity 40.0%; Pred. No. 16;
Matches 6; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 2 EEKPLTTAAAXAPVV 16
|||:|||||
Db 18 QRTTPVSTAGSPVM 32
|||:|||||

RESULT 40
C83492

hypothetical protein PA1228 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: C83492
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
 A:Reference number: AB2950; MUID:20437337; PMID:10984043
 A:Accession: C83492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-116 <STO>
 A:Cross-references: GB:AB004552; GB:AB004091; NID:99947150; PIDN:AAG04617.1; GSPDB:GN001
 A:Experimental source: strain PAO1
 C:Genetics:
 A:Gene: PA1228

Query Match 45.7%; Score 37; DB 2; Length 116;
 Best Local Similarity 50.0%; Pred. No. 28;
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 PLTTAAKAPVXNA 19
 |:|:|:|:
 Db 54 PVTPTAAAPVIDKA 67

Search completed: October 4, 2004, 15:17:53
 Job time : 75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 14:18:24 ; Search time 39 Seconds
(without alignments)
25.368 Million cell updates/sec

Title: US-09-359-426C-2

Perfect score: 81

Sequence: 1 XEEKPLTTAAAPVXNA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	77.8	482	1	CATA_PSEAE
2	53	65.4	484	1	CATA_PROMI
3	45	55.6	128	1	RL7_STRAT
4	43	53.1	483	1	BCA_STRVL
5	41	50.6	482	1	CATA_BORPE
6	40	49.4	722	1	ACM1_DRONE
7	39	48.1	143	1	H2B_AGABE
8	39	48.1	146	1	DMPC_PSEBP
9	39	48.1	527	1	YB64_YEAST
10	39	48.1	560	1	VNUC_INBLE
11	39	48.1	815	1	GRB_MYXIA
12	38	46.9	143	1	GBP_PSESE
13	38	46.9	307	1	PLD_CORPS
14	38	46.9	396	1	ASSY_STRMU
15	38	46.9	479	1	CATA_PSEBU
16	38	46.9	695	1	EFG1_TREPA
17	38	46.9	747	1	YHHA_CAEEL
18	38	46.9	1162	1	FXLB_HUMAN
19	38	46.9	1679	1	YIO9_YEAST
20	37	45.7	366	1	FTZ1_PYRWO
21	37	45.7	396	1	ASSY_STR33
22	37	45.7	396	1	ASSY_STR45
23	37	45.7	442	1	SDC3_MOUSE
24	37	45.7	482	1	CATA_VIBFI
25	37	45.7	498	1	LCYB_CAPAN
26	37	45.7	500	1	LCYB_LYCES
27	37	45.7	500	1	LCYB_TOBAC
28	37	45.7	500	1	PSD3_ANOST
29	37	45.7	511	1	DHAY_YEAST
30	37	45.7	534	1	YAJE_SCHPO
31	37	45.7	662	1	MUC1_XENLA
32	37	45.7	946	1	FTSK_CAMJE
33	37	45.7	1034	1	GCSP_FLAAN

34 37 45.7 1064 1 ISKS_HUMAN
35 37 45.7 1567 1 ICEN_XANCT
36 37 45.7 3390 1 POLG_DEN3
37 36.5 45.1 235 1 CAVI_CAEEL
38 36 44.4 158 1 HUNB_DROMM
39 36 44.4 214 1 DAMX_SERMA
40 36 44.4 228 1 BLOW_STAEP
41 36 44.4 322 1 ANT_EPI1
42 36 44.4 340 1 CNFH_SCHPO
43 36 44.4 501 1 LCYB_ARATH
44 36 44.4 553 1 SPAK_RAT
45 36 44.4 567 1 ODP2_HAEIN

ALIGNMENTS

RESULT 1
CATA_PSEAE
ID CATA_PSEAE STANDARD; PRT; 482 AA.
AC 052752;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Catalase (EC 1.11.1.6).
GN KATA OR PA4236.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRD1;
RX MEDLINE=99296583; PubMed=10368148;
RA Ma J.-F., Ochsen U.A., Klotz M.G., Nanavakkara V.K., Howell M.L.,
Johnson Z., Posey J.E., Vasil M.L., Monaco J.J., Hassett D.J.;
RT "Bactericferitin A modulates catalase A (Kata) activity and
resistance to hydrogen peroxide in Pseudomonas aeruginosa.";
RL J. Bacteriol. 181:3730-3742(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltz L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Raizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -I- FUNCTION: Decomposes hydrogen peroxide into water and oxygen;
serves to protect cells from the toxic effects of hydrogen
peroxide.
CC -I- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -I- COFACTOR: Heme group.
CC -I- ENZYME REGULATION: BY PEROXIDE AND BPR-BOUND IRON.
CC -I- SUBUNIT: HETEROMULTIMER. POSSIBLY AN ALPHA(2)BETA-HETEROTRIMER
WHERE THE ALPHA SUBUNIT IS A 56 kDa PROTEIN AND THE BETA SUBUNIT A
45 kDa PROTEIN.
CC -I- SIMILARITY: Belongs to the catalase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF047025; AAC03118.1; -;
CC EMBL; AF004841; AAC07624.1; -;
DR

DR PIR; B83113; B83113.
 DR HSP; P42321; 2CAE.
 DR InterPro; IPR002226; Catalase.
 DR Pfam; PF00199; catalase; 1.
 DR PRINTS; PR00067; CATALASE.
 DR ProDom; PD000510; Catalase; 1.
 DR PROSITE; PS00437; CATALASE_1; 1.
 DR PROSITE; PS00438; CATALASE_2; 1.
 DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
 KW Complete proteome.
 FT ACT_SITE 55 55 BY SIMILARITY.
 FT ACT_SITE 128 128 BY SIMILARITY.
 FT METAL 338 338 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 482 AA; 55589 MW; 84E5ABA647CAB414 CRC64;
 Query Match 77.8%; Score 63; DB 1; Length 482;
 Best Local Similarity 82.4%; Pred. No. 0.0024;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 EKETLTTAAAPVYVN 18
 DB 2 EKETLTTAAAPVVDN 18
 RESULT 2
 CATA PROMI STANDARD; PRT; 484 AA.
 ID CATA PROMI STANDARD; PRT; 484 AA.
 AC P42321;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 GN Catalase (EC 1.11.1.6).
 GE KATA.
 DE Proteus mirabilis.
 OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Proteus.
 OX NCBI_TaxID=584;
 RN [1]
 RS SEQUENCE, AND SEQUENCE OF 1-305 FROM N.A.
 RC STRAIN=PR;
 RX MEDLINE=95305957; PubMed=7786407;
 RA Buzy A., Bracchi V., Steriades R., Chroboczek J., Thibault P.,
 RA Gagnon J., Jouve H.-M., Hudry-Clergeon G.;
 RA "Complete amino acid sequence of Proteus mirabilis PR catalase.
 RT Occurrence of a methionine sulfone in the close proximity of the
 RT active site.";
 RL J. Protein Chem. 14:59-72(1995).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RC STRAIN=PR;
 RX MEDLINE=95311317; PubMed=7791219;
 RA Guet P., Jouve H.-M., Dideberg O.;
 RA "Crystal structure of Proteus mirabilis PR catalase with and without
 bound NADPH.";
 RL J. Mol. Biol. 249:933-954(1995).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS), AND ABSORPTION SPECTROSCOPY.
 RX MEDLINE=97057535; PubMed=8901874;
 RA Guet P., Jouve H.-M., Williams P.A., Andersson I., Andreoletti P.,
 RA Nussaume L., Hajdu J.;
 RA "Ferryl intermediates of catalase captured by time-resolved
 RT Weissberg crystallography and UV-VIS spectroscopy.";
 RL Nat. Struct. Biol. 3:951-956(1996).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND EPR SPECTROSCOPY.
 RX MEDLINE=22374619; PubMed=12486720;
 RA Andreoletti P., Sainz G., Jaquined M., Gagnon J., Jouve H.-M.;
 RA "High-resolution structure and biochemical properties of a recombinant
 RT Proteus mirabilis catalase depleted in iron.";
 RL Proteins 50:261-271(2003).
 CC -!- FUNCTION: Decomposes hydrogen peroxide into water and oxygen;
 serves to protect cells from the toxic effects of hydrogen
 peroxide.
 CC

DR PIR; B83113; B83113.
 DR HSP; P42321; 2CAE.
 DR InterPro; IPR002226; Catalase.
 DR Pfam; PF00199; catalase; 1.
 DR PRINTS; PR00067; CATALASE.
 DR ProDom; PD000510; Catalase; 1.
 DR PROSITE; PS00437; CATALASE_1; 1.
 DR PROSITE; PS00438; CATALASE_2; 1.
 DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
 KW Complete proteome.
 FT ACT_SITE 55 55 BY SIMILARITY.
 FT ACT_SITE 128 128 BY SIMILARITY.
 FT METAL 338 338 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 482 AA; 55589 MW; 84E5ABA647CAB414 CRC64;
 Query Match 77.8%; Score 63; DB 1; Length 482;
 Best Local Similarity 82.4%; Pred. No. 0.0024;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 EKETLTAAAPVVDN 18
 DB 2 EKETLTAAAPVVDN 18
 RESULT 2
 CATA PROMI STANDARD; PRT; 484 AA.
 ID CATA PROMI STANDARD; PRT; 484 AA.
 AC P42321;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 GN Catalase (EC 1.11.1.6).
 GE KATA.
 DN
 OS Proteus mirabilis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Proteus.
 OX NCBI_TaxID=584;
 RN [1]
 RN SEQUENCE, AND SEQUENCE OF 1-305 FROM N.A.
 RC STRAIN=PR;
 RC MEDLINE=95305957; PubMed=7786407;
 RX Buzy A., Bracchi V., Steriades R., Chroboczek J., Thibault P.,
 RA Gagnon J., Jouve H.-M., Hudry-Clergeon G.;
 RA "Complete amino acid sequence of Proteus mirabilis PR catalase.
 RT Occurrence of a methionine sulfone in the close proximity of the
 RT active site.";
 RT J. Protein Chem. 14:59-72(1995).
 RN [2]
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RP STRAIN=PR;
 RC MEDLINE=95311317; PubMed=7791219;
 RX Guet P., Jouve H.-M., Dideberg O.;
 RA "Crystal structure of Proteus mirabilis PR catalase with and without
 bound NADPH.";
 RT J. Mol. Biol. 249:933-954(1995).
 RN [3]
 RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS), AND ABSORPTION SPECTROSCOPY.
 RP MEDLINE=97057535; PubMed=8901874;
 RX Guet P., Jouve H.-M., Williams P.A., Andersson I., Andreoletti P.,
 RA Nussaume L., Hajdu J.;
 RA "Ferryl intermediates of catalase captured by time-resolved
 RT Weissberg crystallography and UV-VIS spectroscopy.";
 RL Nat. Struct. Biol. 3:951-956(1996).
 RN [4]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS), AND EPR SPECTROSCOPY.
 RP MEDLINE=22374619; PubMed=12486720;
 RX Andreoletti P., Sainz G., Jaquinod M., Gagnon J., Jouve H.-M.;
 RA "High-resolution structure and biochemical properties of a recombinant
 RT Proteus mirabilis catalase depleted in iron.";
 RL Proteins 50:261-271(2003).
 CC -!- FUNCTION: Decomposes hydrogen peroxide into water and oxygen;
 serves to protect cells from the toxic effects of hydrogen
 peroxide.
 CC

```

FT TURN 315 316
FT TURN 320 321
FT TURN 322 323
FT STRAND 328 344
FT HELIX 345 348
FT TURN 347 348
FT TURN 349 351
FT HELIX 353 355
FT TURN 369 370
FT TURN 386 387
FT HELIX 394 396
FT HELIX 412 414
FT HELIX 420 426
FT TURN 427 428
FT HELIX 431 445
FT TURN 446 447
FT HELIX 450 463
FT HELIX 465 476
FT TURN 477 477
SQ SEQUENCE 484 AA; 55614 MW; ADC25F3CB41F5C50 CRC64;

Query Match 65.4%; Score 53; DB 1; Length 484;
Best Local Similarity 75.0%; Pred. No. 0.13;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 EKKPLTTAAAPVYN 18
DB 2 EKKXLTAAAGAPVVDN 17

RESULT 3
RL7_STRAT
ID RL7_STRAT STANDARD; PRT; 128 AA.
AC P29342;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
GN RPLU.
OS Streptomyces antibioticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A., PubMed=1511874;
RX MEDLINE=92380478; PubMed=1511874;
RA Parra F., Blanco G., Alonso J.M., Balbin M., Mendez C., Salas J.A.;
RT "Cloning and sequence of a gene encoding the L7/L12 ribosomal protein
RT equivalent of Streptomyces antibioticus.";
RL Gene 118:127-129(1992)
CC -!- FUNCTION: Seems to be the binding site for several of the factors
CC involved in protein synthesis and appears to be essential for
CC accurate translation (By similarity).
CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M89911; AAA26811.1; -.
CC PIR; JCI1273; JCI1273.
CC HSSP; P02392; ICTP.
CC HAMAP; MF_00368; -.
CC InterPro; IPR008932; Ribos_L12/7_olig.
CC InterPro; IPR00206; Ribosomal_L12.
CC Pfam; PF00542; Ribosomal_L12; 1.
CC ProDom; PD001326; Ribosomal_L12; 1.
CC TIGRFAMs; TIGR00855; L12; 1.
CC Ribosomal protein.

```

```

SQ SEQUENCE 128 AA; 13272 MW; F5C3EE4F45D606E8 CRC64;

Query Match 55.6%; Score 45; DB 1; Length 128;
Best Local Similarity 66.7%; Pred. No. 0.79;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 EKKPLTTAAAPVYN 16
DB 28 EKKFDVTAAPVYN 42

RESULT 4
BCA_STRVL
ID BCA_STRVL STANDARD; PRT; 483 AA.
AC P33569;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bromoperoxidase-catalase (EC 1.11.1.1.-).
GN BCA.
OS Streptomyces violaceus (Streptomyces venezuelae).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1936;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=ATCC 10712 / ISP5230;
RX MEDLINE=97022081; PubMed=8868441;
RA Facey S., Gross F., Vining L.C., Yang K., van Pee K.-H.;
RT "Cloning, sequencing and disruption of a bromoperoxidase-catalase
RT gene in Streptomyces venezuelae: evidence that it is not required for
RT chlorination in chloramphenicol biosynthesis.";
RL Microbiology 142:657-665(1996).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- SIMILARITY: Belongs to the catalase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74791; CAA52796.1; -.
CC PIR; S37055; S37055.
CC HSSP; P42321; 2CAB.
CC InterPro; IPR002236; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom; PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
KW ACT SITE 54 54 BY SIMILARITY.
FT ACT SITE 127 127 BY SIMILARITY.
FT METAL 337 337 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 483 AA; 54087 MW; B58CF8230B8A2F55 CRC64;

Query Match 53.1%; Score 43; DB 1; Length 483;
Best Local Similarity 69.2%; Pred. No. 6.8;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 PLTTAAAPVYN 18
DB 5 PLTTAAGAPVADN 17

RESULT 5
CATA_BORPE
ID CATA_BORPE STANDARD; PRT; 482 AA.
AC P48052;
DT 01-FEB-1996 (Rel. 33, Created)

```


FT DOMAIN 163 184 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 185 208 5 (POTENTIAL).
FT DOMAIN 209 634 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 635 656 6 (POTENTIAL).
FT DOMAIN 657 672 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 673 693 7 (POTENTIAL).
FT DOMAIN 694 722 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 7 7 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 75 75 T -> A (IN REF. 2).
FT CONFLICT 119 119 N -> S (IN REF. 2).
FT CONFLICT 136 136 G -> R (IN REF. 2).
FT CONFLICT 147 147 A -> AA (IN REF. 2).
FT CONFLICT 250 250 P -> G (IN REF. 2).
FT CONFLICT 346 362 MISSING (IN REF. 2).
FT CONFLICT 381 381 G -> A (IN REF. 2).
FT CONFLICT 451 451 T -> A (IN REF. 2).
FT CONFLICT 451 451 C -> VG (IN REF. 2).
FT CONFLICT 606 606 P -> AR (IN REF. 2).
FT CONFLICT 614 614 VLI -> CXS (IN REF. 1).
FT CONFLICT 654 656 S -> C (IN REF. 2).
FT CONFLICT 688 688 EGMVGVN -> DFVAASIR (IN REF. 1).
FT CONFLICT 714 722 EGMVGVN -> DFVAASIR (IN REF. 1).
SQ SEQUENCE 722 AA; 78237 MW; FB35417CB220202A61 CRC64;
Query Match 49.4%; Score 40; DB 1; Length 722;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
Qy 3 EKTPLTTAAAPVXNA 19
Db 538 EDGPTTAAAPLASAA 554
RESULT 7
ID H2B AGABI STANDARD; PRT; 143 AA.
AC P78567;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Histone H2B.
GN HTBA.
OS Agaricus bisporus (Common mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Agaricaceae; Agaricus.
OX NCBI_TaxID=5341;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=971111990; PubMed=8953726;
RA Sonnenberg A.S.M.; de Groot P.W.J.; Schaap P.J.; Baars J.J.P.,
RA Visser J.; van Griensven L.J.L.D.;
RT "Isolation of expressed sequence tags of Agaricus bisporus and their assignment to chromosomes."
RL Appl. Environ. Microbiol. 62:4542-4547(1996).
CC -1- SUBUNIT: The nucleosome is an octamer containing two molecules each of H2A, H2B, H3 and H4. The octamer wraps approximately 146 bp of DNA.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: Belongs to the histone H2B family.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
EMBL; X94188; CAA63898.1;
DR InterPro; IPR007124; Hist_TAF.
DR InterPro; IPR007125; Histone_core_D.
DR InterPro; IPR00558; Histone_H2B.

DR Pfam; PF00125; histone; 1.
DR PRINTS; PR00621; HISTONEH2B.
DR ProDom; PD000497; Histone_H2B; 1.
DR SMART; SM00427; H2B; 1.
DR PROSITE; PS00357; HISTONE_H2B; 1.
KW Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding.
SQ SEQUENCE 143 AA; 15165 MW; 69640102E31F4B56 CRC64;
Query Match 48.1%; Score 39; DB 1; Length 143;
Best Local Similarity 50.0%; Pred. No. 9.6;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 4 KPTLTAAAPVXNA 19
Db 11 KAPASTASKAPVKSDA 26
RESULT 8
ID DMPC_PSESP STANDARD; PRT; 486 AA.
AC P19059;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2-hydroxymuconic semialdehyde dehydrogenase (EC 1.2.1.-) (HMSD).
GN DMPC.
OS Pseudomonas sp. (strain CF600).
OC Plasmid pVil50.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90304229; PubMed=2194577;
RA Nordlund I.; Shingler V.;
RT "Nucleotide sequences of the meta-cleavage pathway enzymes 2-hydroxymuconic semialdehyde dehydrogenase and 2-hydroxymuconic semialdehyde hydrolase from Pseudomonas CF600."
RL Biochim. Biophys. Acta 1049:227-230(1990).
CC -1- FUNCTION: 2-HYDROXYMUCONIC ACID SEMIALDEHYDE CAN BE CONVERTED TO 2-HYDROXPENT-2,4-DIENOATE EITHER DIRECTLY BY THE ACTION OF THREE SEQUENTIAL ENZYMES, THE FIRST OF WHICH IS HMSD.
CC -1- PATHWAY: Meta-cleavage pathway for the degradation of phenols, cresols and catechol. Phenol metabolism.
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
EMBL; X52805; CAA36992.1;
DR HSP; P20000; 1AG8.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD; Plasmid.
FT ACT_SITE 254 254 BY SIMILARITY.
FT ACT_SITE 288 288 BY SIMILARITY.
SQ SEQUENCE 486 AA; 51682 MW; F906FCA64185AA68 CRC64;
Query Match 48.1%; Score 39; DB 1; Length 486;
Best Local Similarity 52.9%; Pred. No. 34;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Qy 3 EKTPLTTAAAPVXNA 19
Db 181 EETPLTTALLGEVQA 197

```

RESULT 9
ID_YB64_YEAST STANDARD; PRT; 527 AA.
AC P38314.
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 57.2 kDa protein in MET8-HPC2 intergenic region.
GN YBR214W OR YBR1501.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Rieger M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 114-527 FROM N.A.
RC STRAIN=S288C;
RA Dubois E., el Bakkoury M., Glandsorff N., Messenguy F., Pierard A.,
RA Scherens B., Vierendeels F.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YGL056C AND S.POMBE SDS23.
CC -!- SIMILARITY: Contains 2 CBS domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z36083; CAA85178.1; -.
CC PIR; S46088; S46088.
CC GeneOnline; 138757; -.
CC SGD; S0000418; SDS24.
CC GO; GO:0005727; Cytoplasm; IDA.
CC GO; GO:0006259; P:DNA metabolism; IGI.
CC GO; GO:0006937; P:endocytosis; IMP.
CC GO; GO:0007126; P:meiosis; IEP.
CC InterPro; IPR000644; CBS_domain.
CC Pfam; PF00571; CBS; 4.
CC SMART; SMO0116; CBS; 2.
CC Hypothetical protein; Repeat; CBS domain.
FT DOMAIN 196 251 CBS 1.
FT DOMAIN 283 335 CBS 2.
SQ SEQUENCE 527 AA; 57187 MW; DC2741550A69C154 CRC64;

Query Match 48.1%; Score 39; DB 1; Length 527;
Best Local Similarity 61.5%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKTPLTTAXAP 14
||| ||| |||
DB 52 EESTPPTATAAP 64

RESULT 10
VNUC_INBLE
ID_VNUC_INBLE STANDARD; PRT; 560 AA.
AC P04665;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE Nucleoprotein.
GN NP.
OS Influenza B virus (strain B/Lee/40).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenzavirus B.

```

```

NCBI_TaxID=11535;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=84174071; PubMed=6324462;
RA Briedis D.J., Tobin M.;
RT Influenza B virus genome: complete nucleotide sequence of the
RT Influenza B/lee/40 virus genome RNA segment 5 encoding the
RT nucleoprotein and comparison with the B/Singapore/222/79
RT nucleoprotein.;
RL Virology 133:448-455(1984).
CC -!- FUNCTION: ENCAPSIDATES THE NEGATIVE STRAND VIRAL RNA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K01395; AAA43689.1; -.
CC InterPro; IPR002141; Flu_NP.
CC DR Pfam; PF00506; Flu_NP; 1.
CC KW Nucleoprotein.
CC SEQUENCE 560 AA; 61770 MW; 74C4FEAF9E75A695 CRC64;

Query Match 48.1%; Score 39; DB 1; Length 560;
Best Local Similarity 44.4%; Pred. No. 39;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 2 EKTPLTTAXAPVYXNA 19
||| ||| ||| |||
DB 19 EELTFGTSGATRPPIKPA 36

RESULT 11
GYRB_MYXXA
ID_GYRB_MYXXA STANDARD; PRT; 815 AA.
AC Q33367;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA gyrase subunit B (EC 5.99.1.3).
GN GYRB.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytophacteria; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=ER-15;
RX MEDLINE=98304088; PubMed=9639935;
RA Paitan Y., Boulton N., Ron E.Z., Rosenberg E., Orr E.;
RT "Molecular analysis of the DNA gyrase gene from Myxococcus xanthus.";
RL Microbiology 144:1641-1647(1998).
CC -!- FUNCTION: DNA gyrase negatively supercoils closed circular double-
CC stranded DNA in an ATP-dependent manner and also catalyzes the
CC interconversion of other topological isomers of double-stranded
CC DNA rings, including catenanes and knotted rings.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -!- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
CC breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
CC enzyme forms an A2B2 tetramer.
CC -!- SIMILARITY: Belongs to the type II topoisomerase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```



```

CC  EMBL; AJ000543; CAA04176.1; -.
CC  HSP; P06982; IAU6.
CC  InterPro; IPR003594; ATPbind_ATPase.
CC  InterPro; IPR002288; DNA_gyraseB_C.
CC  InterPro; IPR000565; DNA_gyrB.
CC  InterPro; IPR001241; DNA_topoisomII.
CC  InterPro; IPR006171; Toprim_dom.
CC  Pfam; PF00204; DNA_gyrases_1.
CC  Pfam; PF00986; DNA_gyrases_C_1.
CC  Pfam; PF02518; HATPase_C_1.
CC  Pfam; PF01751; Toprim_1.
CC  PRINTS; PR00418; TP12FAMILY.
CC  ProDom; PD149633; DNA_gyrases_C_1.
CC  SMART; SM00387; HATPase_C_1.
CC  SMART; SM00433; TOP2C_1.
CC  TIGRFAMs; TIGR01059; gyrB; 1.
CC  ProSITE; PS00177; TOPOISOMERASE_II; 1.
CC  Topoisomerase; Isomerase; ATP-binding.
CC  KW Topoisomerase; Isomerase; ATP-binding.
CC  SQ SEQUENCE 815 AA; 89636 MW; 3862685FBB805B32 CRC64;

Query Match 48.1%; Score 39; DB 1; Length 815;
Best Local Similarity 66.7%; Pred. NO. 57;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 EKTPLTAAAXAP 14
    ||||| :|||
DB 2 EKTPATGSAVAP 13

RESULT 12
GBP_PSESE
ID GBP_PSESE STANDARD; PRT; 143 AA.
AC Q27313;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE Growth-blocking peptide precursor (GBP).
OS Pseudaletia separata (Amyworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Hadeninae; Pseudaletia.
OX NCBI_TaxID=71105;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96105377; PubMed=7498538;
RA Hayakawa Y., Ohnishi A., Yamanaka A., Izumi S., Tomino S.;
RT "Molecular cloning and characterization of cDNA for insect biogenic
RT peptide, growth-blocking peptide.";
RL FEBS Lett. 376:185-189(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98316655; PubMed=9654083;
RA Hayakawa Y., Noguchi H.;
RT "Growth-blocking peptide expressed in the insect nervous system:
RT cloning and functional characterization.";
RL Eur. J. Biochem. 253:810-816(1998).
RN [3]
RP STRUCTURE BY NMR OF 121-143.
RC TISSUE=Hemolymph;
RA Aizawa T., Fujitani N., Hayakawa Y., Ohnishi A., Ohkubo T., Kumaki Y.,
RA Kawano K., Hixichi K., Nitta K.;
RT "Solution structure of an insect growth factor, growth-blocking
RT peptide.";
RL J. Biol. Chem. 274:1887-1890(1999).
CC -!- FUNCTION: Biogenic peptide that prevents, in lepidopteran, the
CC onset of metamorphosis from larva to pupa. This growth-blocking
CC peptide has repressive activity against juvenile hormone esterase.
CC -!- SIMILARITY: Belongs to the GBP / PSP1 / paralytic peptide family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC  the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC  EMBL; S80564; AAB35742.1; -.
CC  EMBL; AB012294; BAA32793.1; -.
CC  PIR; S68226; S68226.
CC  PDB; 1BOF; 09-DEC-98.
CC  InterPro; IPR003463; GBP_PSP.
CC  Pfam; PF02425; GBP_PSP; 1.
CC  ProDom; PD006507; GBP_PSP; 1.
CC  Hemolymph; Signal; 3D-structure.
KW SIGNAL; 1 22 POTENTIAL.
FT PROPEP 23 120 GROWTH-BLOCKING PEPTIDE.
FT PEPTIDE 121 143
FT DISULFID 127 139
FT TURN 129 130
FT STRAND 131 133
FT STRAND 139 141
FT SEQUENCE 143 AA; 15256 MW; DEBFB527956840EB CRC64;

Query Match 46.9%; Score 38; DB 1; Length 143;
Best Local Similarity 50.0%; Pred. NO. 14;
Matches 10; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 2 EEKTP--LTAAAXAPVVXNA 19
    ||||| :|||
DB 83 DEVTPTATTTTAAAPVNA 102

RESULT 13
PLD_CORPS
ID PLD_CORPS STANDARD; PRT; 307 AA.
AC P20626; Q59314;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17; Last sequence update)
DT 01-NOV-1997 (Rel. 35; Last annotation update)
DE Phospholipase D precursor (EC 3.1.1.4) (PLD) (Choline phosphatase).
OS PLD.
OC Corynebacterium pseudotuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1719;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-47.
RX MEDLINE=90170833; PubMed=2407718;
RA Hodgson A.L.M., Bird P., Nisbet I.T.;
RT "Cloning, nucleotide sequence, and expression in Escherichia coli of
RT the phospholipase D gene from Corynebacterium pseudotuberculosis.";
RL J. Bacteriol. 172:1256-1261(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Biovar ovis / Isolate Whetten 1;
RX MEDLINE=9003451; PubMed=2403529;
RA Sanger J.G., Libby S.J., Iandolo J.J., Cuevas W.A.;
RT "Cloning and expression of the phospholipase D gene from
RT Corynebacterium pseudotuberculosis in Escherichia coli.";
RL Infect. Immun. 58:131-136(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Biovar equi / Isolate 155;
RX MEDLINE=95255633; PubMed=7737503;
RA McNamara P.J., Cuevas W.A., Songer J.G.;
RT "Toxic phospholipases D of Corynebacterium pseudotuberculosis, C.
RT ulcerans and Arcanobacterium haemolyticum: cloning and sequence
RT homology.";
RL Gene 156:113-118(1995).
RN [4]
RP MUTAGENESIS.
RC STRAIN=Biovar ovis / Isolate Whetten 1;
RX MEDLINE=95020614; PubMed=7934899;

```

RA McNamee P.J., Bradley G.A., Songer J.G.;
 RT "targeted mutagenesis of the phospholipase D gene results in
 RL decreased virulence of *Corynebacterium pseudotuberculosis*,"
 RL Mol. Microbiol. 12:921-930(1994).
 CC -!- FUNCTION: VIRULENCE FACTOR AFFECTING BACTERIAL DISSIMINATION AND
 CC SURVIVAL WITHIN THE HOST. HAS MAGNESIUM-DEPENDENT SPHINGOMYELINASE
 CC AND HEMOLYTIC ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = choline + a
 CC phosphatidate.
 CC -!- SIMILARITY: TO OTHER CORYNEBACTERIUM PHOSPHOLIPASES D.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L16587; AAA64910.1; -.
 CC EMBL; L16586; AAA99867.1; -.
 CC PIR; A35125; A35125.
 KW Hydrolase; Lipid degradation; Magnesium; Virulence; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 307
 FT ACT SITE 44 44 PHOSPHOLIPASE D.
 FT VARIANT 5 6 POTENTIAL.
 FT VARIANT 8 8 VV -> FA (IN BIOVAR EQUI / ISOLATE 155).
 FT VARIANT 189 189 F -> L (IN BIOVAR EQUI / ISOLATE 155).
 FT VARIANT 205 205 E -> G (IN BIOVAR EQUI / ISOLATE 155).
 FT VARIANT 270 270 N -> D (IN BIOVAR EQUI / ISOLATE 155).
 FT VARIANT 277 277 I -> M (IN BIOVAR EQUI / ISOLATE 155).
 FT VARIANT 277 277 A -> P (IN BIOVAR EQUI / ISOLATE 155).
 SQ SEQUENCE 307 AA; 33884 MW; D3B133456FC99875 CRC64;
 CC
 CC Query Match 46.9%; Score 38; DB 1; Length 307;
 CC Best Local Similarity 61.5%; Pred. No. 31;
 CC Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 6 PLTTAAAPVYXN 18
 CC DB 18 PVGNAAAAPVYHN 30
 CC
 CC RESULT 14
 CC ASSY_STRMU STANDARD; PRT; 396 AA.
 CC AC Q8CZW0;
 CC DT 10-OCT-2003 (Rel. 42, Created)
 CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
 CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
 CC DE Argininosuccinate synthase (EC 6.3.4.5) (Citruelline-aspartate
 CC ligase).
 CC GN ARGG OR SMU.334.
 CC OS Streptococcus mutans.
 CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC Streptococcus.
 CC NCBI_TaxID=1309;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=UAI59 / ATCC 700610 / serotype C;
 CC RX MEDLINE=2295063; PubMed=1239718;
 CC RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 CC Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 CC Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 CC "Genome sequence of *Streptococcus mutans* UAI59, a cariogenic dental
 CC pathogen,"
 CC Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 CC -!- CATALYTIC ACTIVITY: ATP + L-citrulline + L-aspartate = AMP +
 CC di-phosphate + L-argininosuccinate.
 CC -!- PATHWAY: Arginine biosynthesis; seventh step.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -!- SIMILARITY: Belongs to the argininosuccinate synthase family.

CC Subfamily 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AEO14881; AAN58093.1; -.
 CC HAMAP; MF_00005; -; 1.
 CC InterPro; IPR001518; Arginosuc synth.
 CC Pfam; PF00764; Arginosuc synth; 1.
 CC ProDom; PD003544; Arginosuc synth; 1.
 CC TIGRPFAMs; TIGR00032; argG; 1.
 CC PROSITE; PS00564; ARGININOSUCCIN SYN 1; 1.
 CC PROSITE; PS00565; ARGININOSUCCIN SYN 2; 1.
 KW Arginine biosynthesis; Ligase; ATP-binding; Complete proteome.
 SQ SEQUENCE 396 AA; 43913 MW; 44A7C6D56865137E CRC64;
 CC
 CC Query Match 46.9%; Score 38; DB 1; Length 396;
 CC Best Local Similarity 46.7%; Pred. No. 40;
 CC Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 CC
 CC QY 2 EKETLTTAAAPV 16
 CC DB 82 EQKPLVLSALSRPVI 96
 CC
 CC RESULT 15
 CC CATA_PSEPU STANDARD; PRT; 479 AA.
 CC AC Q59714;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Catalase (EC 1.11.1.6).
 CC GN KATA OR CATA.
 CC OS Pseudomonas putida.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 CC Pseudomonadaceae; Pseudomonas.
 CC NCBI_TaxID=303;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=Corvallis;
 CC RX MEDLINE=98019091; PubMed=9358059;
 CC RA Kim Y.C., Miller C.D., Anderson A.J.;
 CC "Identification of adjacent genes encoding the major catalase and a
 CC bacterioferritin from the plant-beneficial bacterium *Pseudomonas*
 CC putida,"
 CC Gene 199:219-224(1997).
 CC -!- FUNCTION: Decomposes hydrogen peroxide into water and oxygen;
 CC serves to protect cells from the toxic effects of hydrogen
 CC peroxide.
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
 CC -!- COFACTOR: Heme group.
 CC -!- ENZYME REGULATION: ACTIVATED BY PEROXIDE.
 CC -!- SIMILARITY: Belongs to the catalase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U63511; AAB88219.1; -.
 CC HSP; P42321; 2CAE.
 CC InterPro; IPR002226; Catalase.
 CC Pfam; PF00199; catalase; 1.
 CC PRINTS; PR00067; CATALASE.

```

DR PRODM; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE 1; 1.
DR PROSITE; PS00438; CATALASE 2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
FT ACT_SITE 53 53 BY SIMILARITY.
FT ACT_SITE 126 126 BY SIMILARITY.
FT METAL 336 336 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 479 AA; 53381 MW; EPE3CBDE67778571 CRC64;

Query Match 46.9%; Score 38; DB 1; Length 479;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 LTTAXAPVXN 18
    |||: |||
Db 5 LTTASGAPVDN 16

RESULT 16
EFGI TREPA
ID EFGI TREPA STANDARD; PRT; 695 AA.
AC O83748;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor G 1 (EF-G 1).
GN FUSA OR FUSA-2 OR TP0767.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OC NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA McDonald J., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388(1998).
CC -!- FUNCTION: This protein promotes the GTP-dependent translocation of
CC the nascent protein chain from the A-site to the P-site of the
CC ribosome.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the GTP-binding elongation factor family.
CC EF-G/EF-2 subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE001248; AAC65735.1; --
DR PIR; D71283; D71283.
DR HSSP; P13551; 2EPG.
DR TIGR; TP0767; --
DR HAMAP; MF 00054; -- 1.
DR InterPro; IPR004540; EF-G.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR009022; EFG_III_V.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR005225; Small GTP.
DR InterPro; IPR009000; Translat_factor.
DR Pfam; PF00679; EFG_C; 1.

```

```

DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFT.
DR TIGRams; TIGR00484; EF-G; 1.
DR TIGRams; TIGR00231; small GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Multigene family; Complete proteome.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 82 86 GTP (BY SIMILARITY).
FT NP_BIND 135 139 GTP (BY SIMILARITY).
SQ SEQUENCE 695 AA; 76831 MW; 46529989BF97E4E CRC64;

Query Match 46.9%; Score 38; DB 1; Length 695;
Best Local Similarity 60.0%; Pred. No. 72;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 BEKTLTTAAAPVV 16
    |||: |||
Db 295 EEPVLSTADAPVV 309

RESULT 17
YMHA CAEEL
ID YMHA CAEEL STANDARD; PRT; 747 AA.
AC P34478;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein F58A4.11 in chromosome III.
GN F58A4.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey I., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans";
RL Nature 368:32-38(1994).
CC -!- SIMILARITY: Contains 2 BED-type zinc fingers.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z22173; CAA80127.1; --
DR EMBL; Z22179; CAA80127.1; JOINED.
DR EMBL; Z22179; CAA80171.1; --
DR EMBL; Z22173; CAA80171.1; JOINED.
DR PIR; F88561; F88561.
DR WormPep; F58A4.11; CRO0217.
DR InterPro; IPR003656; BED_finger.
DR PROSITE; PS50808; ZF_BED; 2.
DR Hypothetical protein; Repeat; Zinc-finger.
KW

```

```

FT ZN FING 250 298 BED-TYPE 1.
FT ZN FING 480 529 BED-TYPE 2.
SQ SEQUENCE 747 AA; 83221 MW; 23890C5F9DCFFB15 CRC64;

Query Match 46.9%; Score 39; DB 1; Length 747;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TPLTTAAAXAPV 16
DB 229 SPCITRASAPSV 240

RESULT 18
FXLB HUMAN STANDARD; PRT; 1162 AA.
AC Q9Y2K7; Q9H7H5; Q9UX66.
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE F-box/UBR-repeat protein 11 (F-box and leucine-rich repeat protein 11)
DE (F-box protein FBL7) (F-box protein Lilina).
GN FBXL11 OR FBL7 OR KIAA1004.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20199565; PubMed=10945468;
RA Ilyin G.P., Riialand M., Pigeon C., Guven-Guillouzo C.;
RT "cDNA cloning and expression analysis of new members of the mammalian
RT F-box protein family.";
RL Genomics 67:40-47(2000).
RN [2]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Brain;
RX MEDLINE=99246063; PubMed=10231032;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RL for large proteins in vitro.";
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RL spleen.";
RN [4]
RP SEQUENCE OF 567-1162 FROM N.A.
RA Pagano M.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex
CC (by similarity).
CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in
CC brain, testis and ovary, followed by lung.
CC -!- SIMILARITY: Contains 1 CXXC-type zinc finger.
CC -!- SIMILARITY: Contains 1 F-box domain.
CC -!- SIMILARITY: Contains 3 leucine-rich (LRR) repeats.
CC -!- SIMILARITY: Contains 1 PHD-type zinc finger.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to a
CC frameshift in position 410.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 238059; CAA86129.1; -
CC PIR; S48385; S48385.

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB023221; BAA76848.2; ALT_INIT.
CC EMBL; AK024505; BAB15795.1; ALT_SEQ.
CC EMBL; AF179221; AAD56012.1; ALT_INIT.
CC Genew; HGNC:13606; FBXL11.
CC MIM; 605657; -
CC
CC InterPro; IPR001810; F-box.
CC InterPro; IPR007089; LRR_Cys.
CC InterPro; IPR003347; TF_JmJC.
CC InterPro; IPR002857; ZnF_CXXC.
CC InterPro; IPR001965; ZnF_PHD.
CC Pfam; PF00646; F-box; 1.
CC Pfam; PF02373; JmJC; 1.
CC Pfam; PF02008; ZF-CXXC; 1.
CC SMART; SM00249; PHD; 1.
CC PROSITE; PS00181; FBOX; FALSE_NEG.
CC PROSITE; PS01359; ZF_PHD 1; 1.
CC PROSITE; PS0016; ZF_PHD 2; 1.
CC Ubl conjugation pathway; Zinc-finger; Leucine-rich repeat; Repeat.
FT ZN FING 564 610 CXXC-TYPE.
FT ZN RING 617 678 PHD-TYPE.
FT DOMAIN 889 936 F-BOX.
FT REPEAT 1000 1025 LRR 1.
FT REPEAT 1063 1091 LRR 2.
FT REPEAT 1095 1118 LRR 3.
SQ SEQUENCE 1162 AA; 132792 MW; 88620A363A5C5842 CRC64;

Query Match 46.9%; Score 38; DB 1; Length 1162;
Best Local Similarity 46.7%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 TPLTTAAAXAPV 19
DB 550 TPVRPAAASPIVSGA 564

RESULT 19
YIO9 YEAST STANDARD; PRT; 1679 AA.
AC P40457;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical 195.1 kDa protein in DNA43-UBI1 intergenic region.
GN YIL149C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97311266; PubMed=9169870;
RA Churcher C.M., Bowman S., Badcock K., Bankier A., Brown D.,
RA Chillingworth T., Connor R., Devlin K., Gentles S., Hamlin N.,
RA Harris D.E., Horneill T., Hunt S., Jagels K., Jones M., Lye G.,
RA Moulis S., Odell C., Pearson D., Rajandream M.A., Rice P., Rowley N.,
RA Skelton J., Smith V., Walsh S., Whitehead S., Barrell B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome IX.";
RL Nature 387:84-87(1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 238059; CAA86129.1; -
CC PIR; S48385; S48385.

```

```

DR  GermOnline; 139684; -
DR  SGD; S0001411; MLP2.
DR  GO; GO:0005635; C:nuclear membrane; IDA.
DR  GO; GO:0005634; C:nucleoplasm; IDA.
DR  GO; GO:0006066; P:protein-nucleus import; IDA.
KW  Hypothetical protein.
SQ  SEQUENCE 1679 AA; 195141 MW; 298950CC52202D8F CRC64;

Query Match      46.9%; Score 38; DB 1; Length 1679;
Best Local Similarity 46.7%; Pred. No. 1.8e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY  2 EKTPLTTAAXAPVV 16
    ||:|||||:|
Db  346 EEEATMTTSAVSPTV 360

RESULT 20
FTZL_PYRWO
ID  FTZL_PYRWO STANDARD; PRT; 366 AA.
AC  Q52630;
DT  01-NOV-1997 (Rel. 35; Created)
DT  01-NOV-1997 (Rel. 35; Last sequence update)
DT  28-FEB-2003 (Rel. 41; Last annotation update)
DE  Cell division protein ftzS homolog 1.
GN  FTZS1 OR FTZS.
OS  Pyrococcus woessli.
OS  Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC  Pyrococcus.
OC  NCBI.
OX  NCBI_TaxID=2262;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=96270610; PubMed=8692886;
RA  Baumann P., Jackson S.P.;
RT  "An archaeobacterial homologue of the essential eubacterial cell
    division protein ftzS."
PL  Proc. Natl. Acad. Sci. U.S.A. 93:6726-6730(1996).
CC  -!- FUNCTION: This protein is essential to the cell-division process.
    It seems to assemble into a dynamic ring on the inner surface of
    the cytoplasmic membrane at the place where division will occur,
    and the formation of the ring is the signal for separation to
    begin. Binds to and hydrolyzes GTP (By similarity).
CC  -!- SUBUNIT: Aggregates to form a ring-like structure (By similarity).
CC  -!- SIMILARITY: Belongs to the ftzS family.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR  EMBL; U56247; AAA99162.1; -.
DR  HSSP; Q57816; 1FGZ.
DR  InterPro; IPR000158; FtsZ.
DR  InterPro; IPR008280; Tub FtsZ C.
DR  InterPro; IPR003008; Tubulin_FtsZ.
DR  Pfam; PF00091; tubulin; 1.
DR  Pfam; PF03953; tubulin_C; 1.
DR  PRINTS; PR00423; CELLDIVISFSTS.
DR  TIGRFAMs; TIGR00065; ftsZ; 1.
DR  PROSITE; PS01134; FTSZ_1; 1.
DR  PROSITE; PS01135; FTSZ_2; 1.
KW  Cell division; Separation; GTP-binding; Multigene family.
FT  NP_BIND 128 136 GTP (POTENTIAL).
SQ  SEQUENCE 366 AA; 39511 MW; 8B40594318F95BFF CRC64;

Query Match      45.7%; Score 37; DB 1; Length 366;
Best Local Similarity 46.7%; Pred. No. 56;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY  2 EKTPLTTAAXAPVV 16
    ||:|||||:|
Db  346 EEEATMTTSAVSPTV 360

RESULT 21
ASSY_STRAS
ID  ASSY_STRAS STANDARD; PRT; 396 AA.
AC  Q8E7N1;
DT  10-OCT-2003 (Rel. 42; Created)
DT  10-OCT-2003 (Rel. 42; Last sequence update)
DT  10-OCT-2003 (Rel. 42; Last annotation update)
DE  Argininosuccinate synthase (EC 6.3.4.5) (Citrulline--aspartate
    ligase).
GN  ARGG OR GBS0123.
OS  Streptococcus agalactiae (serotype III).
OS  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC  Streptococcus.
OC  NCBI.
OX  NCBI_TaxID=216495;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=NEM316 / Serotype III;
    MEDLINE=224242508; PubMed=12354221;
RA  Glaser P., Rusniok C., Buchrieser C., Chevallier F., Frangeul L.,
    Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
    Kunst F.;
RT  "Genome sequence of Streptococcus agalactiae, a pathogen causing
    invasive neonatal disease."
RL  Mol. Microbiol. 45:1499-1513(2002).
CC  -!- CATALYTIC ACTIVITY: ATP + L-citrulline + L-aspartate = AMP +
    diphosphate + L-argininosuccinate.
CC  -!- PATHWAY: Arginine biosynthesis; seventh step.
CC  -!- SUBUNIT: Homotetramer (By similarity).
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC  -!- SIMILARITY: Belongs to the argininosuccinate synthase family.
    Subfamily 1.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR  EMBL; AL766843; CAD45768.1; -.
DR  HAMAP; MF_00005; -; 1.
DR  InterPro; IPR001518; Arginosuc synth.
DR  Pfam; PF00764; Arginosuc synth; 1.
DR  ProDom; PD003544; Arginosuc synth; 1.
DR  TIGRFAMs; TIGR00032; argG; 1.
DR  PROSITE; PS00584; ARGININOSUCCIN SYN_1; 1.
DR  PROSITE; PS00585; ARGININOSUCCIN SYN_2; 1.
KW  Arginine biosynthesis; Ligase; ATP-binding; Complete proteome.
SQ  SEQUENCE 396 AA; 43725 MW; 31DB024B0785AC52 CRC64;

Query Match      45.7%; Score 37; DB 1; Length 396;
Best Local Similarity 40.0%; Pred. No. 60;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY  2 EKTPLTTAAXAPVV 16
    ||:|||||:|
Db  82 EQKYPVLSALSRPII 96

RESULT 22
ASSY_STRAS
ID  ASSY_STRAS STANDARD; PRT; 396 AA.
AC  Q8E272;
DT  10-OCT-2003 (Rel. 42; Created)
DT  10-OCT-2003 (Rel. 42; Last sequence update)
DT  10-OCT-2003 (Rel. 42; Last annotation update)
DE  Argininosuccinate synthase (EC 6.3.4.5) (Citrulline--aspartate
    ligase).
GN  ARGG OR GBS0123.
OS  Streptococcus agalactiae (serotype III).
OS  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC  Streptococcus.
OC  NCBI.
OX  NCBI_TaxID=216495;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=NEM316 / Serotype III;
    MEDLINE=224242508; PubMed=12354221;
RA  Glaser P., Rusniok C., Buchrieser C., Chevallier F., Frangeul L.,
    Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
    Kunst F.;
RT  "Genome sequence of Streptococcus agalactiae, a pathogen causing
    invasive neonatal disease."
RL  Mol. Microbiol. 45:1499-1513(2002).
CC  -!- CATALYTIC ACTIVITY: ATP + L-citrulline + L-aspartate = AMP +
    diphosphate + L-argininosuccinate.
CC  -!- PATHWAY: Arginine biosynthesis; seventh step.
CC  -!- SUBUNIT: Homotetramer (By similarity).
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC  -!- SIMILARITY: Belongs to the argininosuccinate synthase family.
    Subfamily 1.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR  EMBL; AL766843; CAD45768.1; -.
DR  HAMAP; MF_00005; -; 1.
DR  InterPro; IPR001518; Arginosuc synth.
DR  Pfam; PF00764; Arginosuc synth; 1.
DR  ProDom; PD003544; Arginosuc synth; 1.
DR  TIGRFAMs; TIGR00032; argG; 1.
DR  PROSITE; PS00584; ARGININOSUCCIN SYN_1; 1.
DR  PROSITE; PS00585; ARGININOSUCCIN SYN_2; 1.
KW  Arginine biosynthesis; Ligase; ATP-binding; Complete proteome.
SQ  SEQUENCE 396 AA; 43725 MW; 31DB024B0785AC52 CRC64;

Query Match      45.7%; Score 37; DB 1; Length 396;
Best Local Similarity 40.0%; Pred. No. 60;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY  2 EKTPLTTAAXAPVV 16
    ||:|||||:|
Db  82 EQKYPVLSALSRPII 96

RESULT 22
ASSY_STRAS
ID  ASSY_STRAS STANDARD; PRT; 396 AA.
AC  Q8E272;
DT  10-OCT-2003 (Rel. 42; Created)
DT  10-OCT-2003 (Rel. 42; Last sequence update)
DT  10-OCT-2003 (Rel. 42; Last annotation update)
DE  Argininosuccinate synthase (EC 6.3.4.5) (Citrulline--aspartate
    ligase).
GN  ARGG OR GBS0123.
OS  Streptococcus agalactiae (serotype III).
OS  Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC  Streptococcus.
OC  NCBI.
OX  NCBI_TaxID=216495;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=NEM316 / Serotype III;
    MEDLINE=224242508; PubMed=12354221;
RA  Glaser P., Rusniok C., Buchrieser C., Chevallier F., Frangeul L.,
    Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
    Kunst F.;
RT  "Genome sequence of Streptococcus agalactiae, a pathogen causing
    invasive neonatal disease."
RL  Mol. Microbiol. 45:1499-1513(2002).
CC  -!- CATALYTIC ACTIVITY: ATP + L-citrulline + L-aspartate = AMP +
    diphosphate + L-argininosuccinate.
CC  -!- PATHWAY: Arginine biosynthesis; seventh step.
CC  -!- SUBUNIT: Homotetramer (By similarity).
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC  -!- SIMILARITY: Belongs to the argininosuccinate synthase family.
    Subfamily 1.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR  EMBL; AL766843; CAD45768.1; -.
DR  HAMAP; MF_00005; -; 1.
DR  InterPro; IPR001518; Arginosuc synth.
DR  Pfam; PF00764; Arginosuc synth; 1.
DR  ProDom; PD003544; Arginosuc synth; 1.
DR  TIGRFAMs; TIGR00032; argG; 1.
DR  PROSITE; PS00584; ARGININOSUCCIN SYN_1; 1.
DR  PROSITE; PS00585; ARGININOSUCCIN SYN_2; 1.
KW  Arginine biosynthesis; Ligase; ATP-binding; Complete proteome.
SQ  SEQUENCE 396 AA; 43725 MW; 31DB024B0785AC52 CRC64;

Query Match      45.7%; Score 37; DB 1; Length 396;
Best Local Similarity 40.0%; Pred. No. 60;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY  2 EKTPLTTAAXAPVV 16
    ||:|||||:|
Db  82 EQKYPVLSALSRPII 96

```

DE ligase).
GN ARGG OR SAG0125.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=2222988; PubMed=12200547;
RA Tettelin H., Masiognani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Kessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouli H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni P., Malone D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Praser C.M.,
RA "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
CC -!- CATALYTIC ACTIVITY: ATP + L-citrulline + L-aspartate = AMP +
CC diphosphate + L-argininosuccinate.
CC -!- PATHWAY: Arginine biosynthesis; seventh step.
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the argininosuccinate synthase family.
CC Subfamily 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE014197; AAM99033.1; -.
DR TIGR; SAG0125; -.
DR HANAP; MF_000055; -; 1.
DR InterPro; IPR001518; Argininosuc synth.
DR Pfam; PF00764; Argininosuc synth; 1.
DR ProDom; PD003544; Argininosuc synth; 1.
DR TIGRFAMs; TIGR00032; argG; 1.
DR PROSITE; PS00564; ARGININOSUCCIN SYN 1; 1.
DR PROSITE; PS00565; ARGININOSUCCIN SYN 2; 1.
DR Arginine biosynthesis; Ligase; ATP-binding; Complete proteome.
KW Arginine biosynthesis; Ligase; ATP-binding; Complete proteome.
SQ SEQUENCE 396 AA; 43697 MW; 31C8C841B6E5AC52 CRC64;

Query Match 45.7%; Score 37; DB 1; Length 396;
Best Local Similarity 40.0%; Pred. No. 60;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 BEKTLPLTAXAPV 16
Db 82 EQKPLVLSALSRPII 96

RESULT 23
SDC3 MOUSE
ID -SDC3_MOUSE STANDARD; PRT; 442 AA.
AC Q64519;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SynDecan-3 precursor (SYND3).
GN SDC3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6 X CBA;
RA Kung C.E., Deuel T.F.;
RT "Cloning of rat and mouse syndecan-3 cDNAs";
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cell surface proteoglycan that may bear heparan sulfate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- PTM: O-GLYCOSYLATED WITHIN THE THR/SER-RICH REGION WHICH COULD
CC INTERACT WITH LECTIN DOMAINS ON OTHER MOLECULES (PROBABLE).
CC -!- SIMILARITY: Belongs to the syndecan proteoglycan family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U52826; AAS03283.1; -.
DR MGD; MGI:1349163; Sdc3.
DR InterPro; IPR003585; Neurexin-like.
DR Pfam; PF01034; Syndecan; 1.
DR SMART; SM00294; 4.lm; 1.
DR PROSITE; PS00564; SYNDECAN; 1.
KW Proteoglycan; Heparan sulfate; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 45
FT CHAIN 46 442
FT DOMAIN 46 384
FT TRANSMEM 385 409
FT DOMAIN 410 442
FT DOMAIN 114 302
FT SITE 383 384
FT CARBOHYD 78 78
FT CARBOHYD 80 80
FT CARBOHYD 82 82
FT CARBOHYD 89 89
FT CARBOHYD 315 315
FT CARBOHYD 367 367
SQ SEQUENCE 442 AA; 46101 MW; D2F1CE6E611FA164 CRC64;

Query Match 45.7%; Score 37; DB 1; Length 442;
Best Local Similarity 57.1%; Pred. No. 67;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 6 PLTTAAXAPVVKNA 19
Db 216 PLTTAATKITTPA 229

RESULT 24
CATA VIBFI
ID CATA_VIBFI STANDARD; PRT; 482 AA.
AC Q68146;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Catalase (EC 1.11.1.6).
GN KATA.
OS Vibrio fischeri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98215175; PubMed=9555890;
RA Visick K., Ruby E.G.;
RT "The periplasmic, group III catalase of Vibrio fischeri is required
RT for normal symbiotic competence and is induced both by oxidative
RT stress and by approach to stationary phase."
RL J. Bacteriol. 180:2087-2092(1998).

CC -!- FUNCTION: Decomposes hydrogen peroxide into water and oxygen;
 CC serves to protect cells from the toxic effects of hydrogen
 CC peroxide. Could protect calls in nodules which have a high
 CC potential to produce hydrogen peroxide because of the strong
 CC reducing conditions required for nitrogen fixation and the action
 CC of several proteins.
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
 CC -!- COFACTOR: Heme group.
 CC -!- SUBCELLULAR LOCATION: Periplasmic.
 CC -!- INDUCTION: By hydrogen peroxide.
 CC -!- MISCELLANEOUS: EITHER KATA OR KATC IS ABSOLUTELY REQUIRED FOR THE
 CC PROTECTION OF THE NITROGEN FIXATION PROCESS.
 CC -!- SIMILARITY: Belongs to the catalase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF011784; AAC38344.1; --
 CC HSSP; P42321; 2CAE.
 CC InterPro; IPR002226; Catalase.
 CC Pfam; PF00199; catalase; 1.
 CC PRINTS; PR00067; CATALASE.
 CC ProDom; PD000510; Catalase; 1.
 CC PROSITE; PS00437; CATALASE_1; 1.
 CC PROSITE; PS00438; CATALASE_2; 1.
 CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
 CC Periplasmic.
 CC ACT_SITE 53 53 BY SIMILARITY.
 CC ACT_SITE 126 126 BY SIMILARITY.
 CC METAL 336 336 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 CC SEQUENCE 482 AA; 54864 MW; 3808F0DBDE4DCC32 CRC64;
 CC
 CC Query Match 45.7%; Score 37; DB 1; Length 482;
 CC Best Local Similarity 66.7%; Pred. No. 74;
 CC Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 CC
 CC QY 7 LTAAAPVWVN 18
 CC |||||
 CC DB 5 LTAAAGCPVAHN 16
 CC
 CC RESULT 25
 CC LCYB CAPAN
 CC ID LCYB CAPAN STANDARD; PRT; 498 AA.
 CC AC Q43415;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Lycopene beta cyclase, chloroplast precursor (EC 1.14.-.-).
 CC GN LCY1 OR CRTL.
 CC OS Capsicum annuum (Bell pepper).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 CC OC Lamnids; Solanales; Solanaceae; Capsicum.
 CC OX NCBI_TaxID=4072;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=cv. Lamuyo; TISSUE=fruit;
 CC MEDLINE=96045549; PubMed=7550379;
 CC RA Hugueney P., Badillo A., Chen H.C., Klein A., Hirschberg J.,
 CC RA Camara B., Kuntz M.;
 CC RT "Metabolism of cyclic carotenoids: a model for the alteration of this
 CC RT biosynthetic pathway in Capsicum annum chromoplasts.";
 CC RL Plant J. 8:417-424(1995).
 CC -!- FUNCTION: Catalyzes the double cyclization reaction which converts
 CC lycopene to beta-carotene and neurosporene to beta-zeacarotene.
 CC -!- PATHWAY: Carotenoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast; chromoplast.

CC -!- SIMILARITY: Belongs to the lycopene cyclase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X86221; CAA60119.1; --
 CC InterPro; IPR001327; PAD_pyr_redox.
 CC InterPro; IPR008671; Lycopene_cycl.
 CC InterPro; IPR000205; NAD_BS.
 CC InterPro; IPR001100; Pyr_redox.
 CC Pfam; PF05834; Lycopene_cycl; 1.
 CC PRINTS; PR00368; PADPNT.
 CC PRINTS; PR00411; PNDPRTASEI.
 CC Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
 CC Transit peptide.
 CC TRANSIT 1 79 CHLOROPLAST (POTENTIAL).
 CC FT CHAIN 80 498 LYCOPENE BETA CYCLASE.
 CC FT NP_BIND 84 112 NAD (POTENTIAL).
 CC SQ SEQUENCE 498 AA; 55610 MW; 177180CD5745F64F CRC64;
 CC
 CC Query Match 45.7%; Score 37; DB 1; Length 498;
 CC Best Local Similarity 72.7%; Pred. No. 76;
 CC Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC QY 9 TAAAPVWANA 19
 CC |||||
 CC DB 370 TAAAPVWANA 380
 CC
 CC RESULT 26
 CC LCYB LYCES
 CC ID LCYB LYCES STANDARD; PRT; 500 AA.
 CC AC Q43503;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Lycopene beta cyclase, chloroplast precursor (EC 1.14.-.-).
 CC GN LCY1 OR CRTL-1.
 CC OS Lycopersicon esculentum (Tomato).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 CC OC Lamnids; Solanales; Solanaceae; Solanum.
 CC OX NCBI_TaxID=4081;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=cv. VF36; TISSUE=Leaf;
 CC MEDLINE=96434545; PubMed=8837512;
 CC RA Cunningham F.X. Jr., Pogson B., Sun Z., McDonald K.A., Dellapenna D.,
 CC RA Gantt E.;
 CC RT "Functional analysis of the beta and epsilon lycopene cyclase enzymes
 CC RT of Arabidopsis reveals a mechanism for control of cyclic carotenoid
 CC RT formation.";
 CC RL Plant Cell 8:1613-1626(1996).
 CC -!- FUNCTION: Catalyzes the double cyclization reaction which converts
 CC lycopene to beta-carotene and neurosporene to beta-zeacarotene.
 CC -!- PATHWAY: Carotenoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: Belongs to the lycopene cyclase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X86452; CAA60170.1; --

```
DR PIR; S72505; S72505.
DR InterPro; IPR001327; FAD pyr_redox.
DR InterPro; IPR008671; Lycopene_cycl.
DR InterPro; IPR000205; NAD BS.
DR InterPro; IPR001100; Pyr_redox.
DR Pfam; PF05834; Lycopene_cycl; 1.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNDRDTASE1.
DR Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
KW Transit peptide.
FT CHAIN 1 81 CHLOROPLAST (POTENTIAL).
FT CHAIN 82 500 LYCOPENE BETA CYCLASE.
FT NP_BIND 86 114 NAD (POTENTIAL).
SQ SEQUENCE 500 AA; 56180 MW; CF42FD7D4684C04DD CRC64;

Query Match 45.7%; Score 37; DB 1; Length 500;
Best Local Similarity 72.7%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 TAAXAPVVXNA 19
DB 372 TLAAAPVVANA 382

RESULT 27
LCYB_TOBAC
ID LCYB_TOBAC STANDARD; PRT; 500 AA.
AC Q43578;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lycopene beta cyclase, chloroplast precursor (EC 1.14.-.-).
CN LCVI OR CRTL-1.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ecv. Samsun NN; TISSUE=Leaf;
RX MEDLINE=96434545; PubMed=8837512;
RA Cunningham F.X. Jr., Pogson B., Sun Z., McDonald K.A., Dellapenna D.,
RA Gantt E.;
RT "Functional analysis of the beta and epsilon lycopene cyclase enzymes
of Arabidopsis reveals a mechanism for control of cyclic carotenoid
formation."
RT Plant Cell 8:1613-1626 (1996).
CC -!- FUNCTION: Catalyzes the double cyclization reaction which converts
lycopene to beta-carotene and neurosporene to beta-zeacarotene.
CC -!- PATHWAY: Carotenoid biosynthesis.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: Belongs to the lycopene cyclase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; X81787; CAA57386.1; -.
DR PIR; S72506; S72506.
DR InterPro; IPR008671; Lycopene_cycl.
DR InterPro; IPR000205; NAD BS.
DR Pfam; PF05834; Lycopene Cycl; 1.
DR Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
KW Transit peptide.
FT CHAIN 1 81 CHLOROPLAST (POTENTIAL).
FT CHAIN 82 500 LYCOPENE BETA CYCLASE.
FT NP_BIND 86 114 NAD (POTENTIAL).
SQ SEQUENCE 500 AA; 56067 MW; 2E3721B87EE6CB8 CRC64;

Query Match 45.7%; Score 37; DB 1; Length 500;
Best Local Similarity 58.3%; Pred. No. 76;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 TPLTTAAXAPVV 16
DB 7 TPTATASEPIV 18

RESULT 29
PSD3_ANOST
ID PSD3_ANOST STANDARD; PRT; 500 AA.
AC Q9USZ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable 26S proteasome non-ATPase regulatory subunit 3 (26S
proteasome subunit S3) (Diphenol oxidase A2 component) (DOX-A2).
CN DOXA2 OR DOX-A2.
OS Anopheles stephensi (Indo-Pakistan malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=30069;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=StMal;
RX MEDLINE=20296811; PubMed=10835480;
RA Garvey C.F., Malcolm C.A.;
RT "Anopheles stephensi Dox-A2 shares common ancestry with genes from
distant groups of eukaryotes encoding a 26S proteasome subunit and is
in a conserved gene cluster."
RL J. Mol. Evol. 50:497-509 (2000).
CC -!- FUNCTION: Acts as a regulatory subunit of the 26 proteasome which
is involved in the ATP-dependent degradation of ubiquitinated
proteins (By similarity).
CC -!- SUBUNIT: The 26S proteasome is composed of a core protease, known
as the 20S proteasome, capped at one or both ends by the 19S
regulatory complex (RC). The RC is composed of at least 18
different subunits in two subcomplexes, the base and the lid,
which form the portions proximal and distal to the 20S proteolytic
core, respectively (By similarity).
CC -!- SIMILARITY: Contains 1 PCI domain.
CC -!- SIMILARITY: Belongs to the proteasome subunit S3 family.
CC -!- CAUTION: Was originally (Ref.1) thought to be the diphenol oxidase
A2 component involved in catecholamine metabolism, melanin
formation, and sclerotization of the cuticle.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; AJ250874; CAB61220.1; -.
DR InterPro; IPR000717; PCI.
DR Pfam; PF01399; PCI; 1.
DR SMART; SM00088; PINT; 1.
KW Proteasome.
SQ SEQUENCE 500 AA; 56840 MW; 4B98EA0F702EB5C CRC64;

Query Match 45.7%; Score 37; DB 1; Length 500;
Best Local Similarity 72.7%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 TAAXAPVVXNA 19
DB 372 TLAAAPVVANA 382
```



```

DHAY YEAST
ID DHAY YEAST STANDARD; PRT; 511 AA.
AC P32872;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aldehyde dehydrogenase 2, mitochondrial precursor (EC 1.2.1.3).
GN ALD2 OR ALDH2.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Thiesen J.;
RL Thiesen (1993), Heinrich-Heine University / Duesseldorf, Germany.
CC -!- CAPALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
CC -!- PATHWAY: Ethanol utilization; second step.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix (Potential).
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
CC -!- CAUTION: Is not present in yeast genome.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z17314; CAA78962.1; -.
DR PIR; S31308; S31308.
DR HSSP; P05091; 1CW3.
DR SGB; L0000075; ALD2.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
DR Oxidoreductase; NAD; Mitochondrion; Transit peptide.
FT TRANSIT 1 21 MITOCHONDRION
FT CHAIN 22 511 ALDEHYDE DEHYDROGENASE 2.
FT NP_BIND 274 279 NAD (ADP PART) (BY SIMILARITY).
FT ACT_SITE 297 297 BY SIMILARITY.
FT ACT_SITE 331 331 BY SIMILARITY.
SQ SEQUENCE 511 AA; 56466 MW; 70EDA951B84EE4A CRC64;
Query Match 45.7%; Score 37; DB 1; Length 511;
Best Local Similarity 44.4%; Pred. No. 78;
Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
QY 2 EKTPLTTAAAPVXNA 19
DB 223 BELTPLTALTVALIKEA 240
RESULT 30
YAJE SCHPO
ID YAJE SCHPO STANDARD; PRT; 534 AA.
AC Q09311.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C30D11.14c in chromosome I.
GN SPAC30D11.14C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA Hauser F.; Hoffmann W.;
RL "P-domains as shuffled cysteine-rich modules in integumentary mucin
RT C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic
RT polymorphism.";

```

```

Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Murgall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weijens J., Vanstreels E., Rieger M., Schaefer M., Muelier-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002)
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z67961; CAA91900.1; -.
DR PIR; T38585; S62572.
DR GenedB SPombe; SPAC30D11.14C; -.
DR InterPro; IPR004087; KH_dom.
DR SMART; SM00322; KH; 1.
KW Hypothetical protein.
SQ SEQUENCE 534 AA; 59035 MW; B38F20B8A20A47F7 CRC64;
Query Match 45.7%; Score 37; DB 1; Length 534;
Best Local Similarity 53.3%; Pred. No. 82;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 EKTPLTTAAAPV 16
DB 415 QEXTPTNAPPLV 429
RESULT 31
MUC1_XENLA
ID MUC1_XENLA STANDARD; PRT; 662 AA.
AC Q05049;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integumentary mucin C.1 (FIM-C.1) (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5; 6 AND 7).
RA Tissue-Skin.
EC MEDLINE=93077556; PubMed=1447205;
RX Hauser F.; Hoffmann W.;
RL "P-domains as shuffled cysteine-rich modules in integumentary mucin
RT C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic
RT polymorphism.";

```


jejun flagellar gene flhA.";
 RL Gene 146:31-38(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parthill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 BA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 RA Jagers K., Kariyeh A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrell B.G.;
 RT "The genome sequence of the food-borne pathogen *Campylobacter jejuni*
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 CC -!- FUNCTION: DNA motor protein, which is both required to move DNA
 CC out of the region of the septum during cell division and for the
 CC septum formation. Tracks DNA in an ATP-dependent manner by
 CC generating positive supercoils in front of it and negative
 CC supercoils behind it (By similarity).
 CC -!- SUBUNIT: Homohexamer. This suggests the formation of a ring
 CC between the two cells at the septum that surrounds DNA (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Located at the
 CC septum. The large C-terminal part of the protein is cytoplasmic
 CC (Potential).
 CC -!- SIMILARITY: Contains 1 FtsK domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U06951; AA061512.1; ALT_INIT.
 DR EMBL; AL139076; CAB73144.1; -;
 DR F1R; F81361; F81361.
 DR HAMAP; MF 01809; -; 1.
 DR InterPro; IPR002543; FtsK_SpoIIIE.
 DR Pfam; PF01580; FtsK_SpoIIIE; 1.
 DR PROSITE; PS05091; FTSK; 1.
 KW Chromosome partition; Cell division; ATP-binding; DNA-binding;
 KW Transmembrane; Complete proteome.
 FT TRANSMEM 15 35 POTENTIAL.
 FT TRANSMEM 50 70 POTENTIAL.
 FT TRANSMEM 86 106 POTENTIAL.
 FT TRANSMEM 521 541 POTENTIAL.
 FT TRANSMEM 731 751 POTENTIAL.
 FT DOMAIN 620 810 FTSK.
 FT NP_BIND 637 644 ATP (POTENTIAL).
 FT CONFLICT 22 22 A -> G (IN REF. 1).
 FT CONFLICT 28 30 FLF -> SLL (IN REF. 1).
 FT CONFLICT 35 35 V -> L (IN REF. 1).
 FT CONFLICT 39 40 RN -> KK (IN REF. 1).
 FT CONFLICT 43 44 LA -> IE (IN REF. 1).
 FT CONFLICT 47 47 T -> K (IN REF. 1).
 FT CONFLICT 54 56 IGF -> FSL (IN REF. 1).
 FT CONFLICT 61 61 S -> T (IN REF. 1).
 FT CONFLICT 66 66 A -> S (IN REF. 1).
 FT CONFLICT 70 72 PMS -> REF (IN REF. 1).
 FT CONFLICT 78 85 LAYAFST -> IVYGFESI (IN REF. 1).
 FT CONFLICT 89 89 H -> R (IN REF. 1).
 FT CONFLICT 103 103 S -> Y (IN REF. 1).
 FT CONFLICT 143 144 EN -> VF (IN REF. 1).
 FT CONFLICT 156 156 V -> A (IN REF. 1).
 FT CONFLICT 378 378 T -> A (IN REF. 1).
 FT CONFLICT 478 478 I -> V (IN REF. 1).
 FT CONFLICT 735 735 K -> E (IN REF. 1).
 FT CONFLICT 888 888 A -> V (IN REF. 1).
 SQ SEQUENCE 946 AA; 107862 MW; 825D15A43BEAB4FD CRC64;
 Query Match 45.7%; Score 37; DB 1; Length 946;
 Best Local Similarity 61.5%; Pred. No. 1.6e+02;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Best Local Similarity 50.0%; Pred. No. 1.5e+02;
 Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 4 KTPLTTAAAPVXXNA 19
 DB 607 KSPLTIALGKDIVGNA 622
 RESULT 33
 GCSP_FLAAN STANDARD; PRT; 1034 AA.
 ID GCSP_FLAAN
 AC 049850;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glycine dehydrogenase [decarboxylating], mitochondrial precursor
 DE (EC 1.4.4.2) (Glycine decarboxylase) (Glycine cleavage system P-
 DE protein).
 GN GCSP.
 OS Flaveria anomala.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC campanulids; Asterales; Asteraceae; Tageteae; Flaveria.
 OC NCB1_TaxID=35877;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RA Nan O., Bauwe H.;
 RT "The GCSP gene encoding P-protein of the glycine cleavage system in
 RT the C3-C4 intermediate plant *Flaveria anomala*.";
 RL (In) Plant Gene Register PGR98-004.
 CC -!- FUNCTION: The glycine cleavage system catalyzes the degradation of
 CC glycine. The P protein binds the alpha-amino group of glycine
 CC through its pyridoxal phosphate cofactor; CO(2) is released and
 CC the remaining methylamine moiety is then transferred to the
 CC lipamide cofactor of the H protein.
 CC -!- CATALYTIC ACTIVITY: Glycine + lipoylprotein = S-
 CC aminomethylidihydropyridoylprotein + CO(2).
 CC -!- COFACTOR: Pyridoxal phosphate.
 CC -!- SUBUNIT: Homodimer (By similarity). The glycine cleavage system is
 CC composed of four proteins: P, T, L and H.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial.
 CC -!- SIMILARITY: Belongs to the gcvp family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z99762; CAB16911.1; -;
 DR InterPro; IPR003437; GDC-P.
 DR Pfam; PF02347; GDC-P; 1.
 DR TIGRPFAMS; TIGR00461; gcvp; 1.
 DR Oxidoreductase; Mitochondrion; Pyridoxal phosphate; Transit peptide.
 KW TRANSIT 1 63 MITOCHONDRION (POTENTIAL)
 FT CHAIN 64 1034 GLYCINE DEHYDROGENASE [DECARBOXYLATING]
 FT BINDING 770 770 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 1034 AA; 112695 MW; C2F424C76EF104BA CRC64;
 Query Match 45.7%; Score 37; DB 1; Length 1034;
 Best Local Similarity 61.5%; Pred. No. 1.6e+02;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 2 EEXTPLTTAAAP 14
 DB 811 EESQPLGTIAAP 823
 RESULT 34
 ISKS_HUMAN

ID ISK5 HUMAN STANDARD; PRT; 1064 AA.
AC Q9N038; O75770; Q96P2; Q96P3;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last sequence update)
DE Serine protease inhibitor Kazal-type 5 precursor (Lympho-epithelial
DE Kazal-type related inhibitor) (LEKTI) [Contains: Hemofiltrate peptide
DE HF6478; Hemofiltrate peptide HF7665].
GN SPINK5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANT GLU-420.
RC TISSUE=Epithelium;
RX MEDLINE=9348267; PubMed=10419450;
RA Maegert H.-J., Staendker L., Kreutzmann P., Zucht H.-D., Reinecke M.,
RA Sommerhoff C.P., Fritz H., Forssmann W.-G.;
RT "LEKTI, a novel 15-domain type of human serine proteinase inhibitor";
RL J. Biol. Chem. 274:21499-21502(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20296612; PubMed=10835624;
RA Chavanas S., Bodemer C., Rochat A., Hamel-Teillac D., Ali M.,
RA Irvine A.D., Bonafe J.-L., Wilkinson J., Taieb A., Barrandon Y.,
RA Harper J.I., de Prost Y., Hovnanian A.;
RT "Mutations in SPINK5, encoding a serine protease inhibitor, cause
RT Netherton syndrome";
RL Nat. Genet. 25:141-142(2000).
RN [3]
RP SEQUENCE OF 202-222 AND 266-294 FROM N.A.
RX MEDLINE=21403066; PubMed=11511292;
RA Sprecher E., Chavanas S., DiGiovanna J.J., Amin S., Nielsen K.,
RA Prendiville J.S., Silverman R., Esterly N.B., Spraker M.K., Guelig E.,
RA de Luna M.L., Williams M.L., Buehler B., Siegfried E.C.,
RA van Maldergem L., Prencher E., Bale S.J., Uitto J., Hovnanian A.,
RA Richard G.;
RT "The spectrum of pathogenic mutations in SPINK5 in 19 families with
RT Netherton syndrome: implications for mutation detection and first case
RT of prenatal diagnosis";
RL J. Invest. Dermatol. 117:179-187(2001).
RN [4]
RP SEQUENCE OF 490-507.
RC TISSUE=Fore skin keratinocyte;
RX MEDLINE=21477652; PubMed=11594460;
RA Ahmed A., Kandola P., Ziada G., Parenteau N.;
RT "Purification and partial amino acid sequence of proteins from human
RT epidermal keratinocyte conditioned medium";
RL J. Protein Chem. 20:273-278(2001).
RN [5]
RP VARIANT GLU-420.
RX MEDLINE=21470331; PubMed=11544479;
RA Walley A.J., Chavanas S., Moffatt M.F., Esnouf R.M., Ubbi B.,
RA Lawrence R., Wong K., Abecasis G.R., Jones E.V., Harper J.I.,
RA Hovnanian A., Cookson W.O.C.M.;
RT "Gene polymorphism in Netherton and common atopic disease";
RL Nat. Genet. 29:175-178(2001).
CC -!- FUNCTION: SERINE PROTEASE INHIBITOR, PROBABLY IMPORTANT FOR THE
CC ANTI-INFLAMMATORY AND/OR ANTIMICROBIAL PROTECTION OF MUCOUS
CC EPITHELIA.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE THYMUS. ALSO FOUND IN
CC THE ORAL MUCOSA, PARATHYROID GLAND, BARTHOLIN'S GLANDS, TONSILS,
CC AND VAGINAL EPITHELIUM. VERY LOW LEVELS ARE DETECTED IN LUNG,
CC KIDNEY, AND PROSTATE.
CC -!- DOMAIN: CONTAINS AT LEAST ONE ACTIVE INHIBITORY DOMAIN FOR TRYPSIN
CC (DOMAIN 6).
CC -!- DISEASE: Defects in SPINK5 are the cause of Netherton syndrome
CC (NS) [MIM:256500], a severe autosomal recessive disorder. It is
CC characterized by congenital ichthyosis with defective
CC cornification, a specific hair shaft defect known as trichorrhexis
CC invaginata or "bamboo hair", atopic dermatitis, and hayfever. High

CC post-natal mortality is due to failure to thrive, infections and
CC hypernatraemic dehydration.
CC -!- SIMILARITY: Contains 15 Kazal-like domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ228139; CAB40839.1; -
CC EMBL; AJ391230; CAB96877.1; -
CC EMBL; AJ270944; CAB96877.1; JOINED.
CC EMBL; AJ391231; CAB96877.1; JOINED.
CC EMBL; AJ391232; CAB96877.1; JOINED.
CC EMBL; AJ391233; CAB96877.1; JOINED.
CC EMBL; AJ391234; CAB96877.1; JOINED.
CC EMBL; AJ391235; CAB96877.1; JOINED.
CC EMBL; AJ276573; CAB96877.1; JOINED.
CC EMBL; AJ391236; CAB96877.1; JOINED.
CC EMBL; AJ276580; CAB96877.1; JOINED.
CC EMBL; AJ391237; CAB96877.1; JOINED.
CC EMBL; AJ391238; CAB96877.1; JOINED.
CC EMBL; AJ391239; CAB96877.1; JOINED.
CC EMBL; AJ391240; CAB96877.1; JOINED.
CC EMBL; AJ391241; CAB96877.1; JOINED.
CC EMBL; AJ276578; CAB96877.1; JOINED.
CC EMBL; AJ391242; CAB96877.1; JOINED.
CC EMBL; AJ391243; CAB96877.1; JOINED.
CC EMBL; AJ391244; CAB96877.1; JOINED.
CC EMBL; AJ391245; CAB96877.1; JOINED.
CC EMBL; AJ391246; CAB96877.1; JOINED.
CC EMBL; AJ391247; CAB96877.1; JOINED.
CC EMBL; AJ391248; CAB96877.1; JOINED.
CC EMBL; AJ391249; CAB96877.1; JOINED.
CC EMBL; AJ391250; CAB96877.1; JOINED.
CC EMBL; AJ391251; CAB96877.1; JOINED.
CC EMBL; AJ391252; CAB96877.1; JOINED.
CC EMBL; AJ391253; CAB96877.1; JOINED.
CC EMBL; AJ391254; CAB96877.1; JOINED.
CC EMBL; AJ276577; CAB96877.1; JOINED.
CC EMBL; AF295784; AAK37139.1; -
CC EMBL; AF295783; AAK37140.1; -
CC HSP; P37109; 1PCE.
CC Genew; HGNC:15464; SPINK5.
CC MIM; 605010; -
CC MIM; 256500; -
CC GO; GO:0005576; C:extracellular; NAS.
CC GO; GO:0004867; F:serine protease inhibitor activity; NAS.
CC GO; GO:0002336; P:anti-inflammatory response; NAS.
CC GO; GO:0006960; P:antimicrobial humoral response (sensu Inver. . . ; NAS.
CC InterPro; IPR002350; Kazal.
CC Pfam; PF00050; Kazal; 7.
CC SMART; SM00280; Kazal; 14.
CC PROSITE; PS00282; KAZAL; 2.
CC Serine protease inhibitor; Repeat; Signal; Polymorphism.
CC SIGNAL 1 22
CC CHAIN 23 1064
CC PEPTIDE 23 77
CC PEPTIDE 356 423
CC DOMAIN 28 66
CC DOMAIN 95 151
CC DOMAIN 159 214
CC DOMAIN 223 272
CC DOMAIN 295 351
CC DOMAIN 365 421
CC DOMAIN 435 473
CC DOMAIN 494 556
CC DOMAIN 565 621
CC DOMAIN 630 686
CC DOMAIN 705 757
CC SERINE PROTEASE INHIBITOR KAZAL-TYPE 5.
CC HEMOFILTRATE PEPTIDE HF6478.
CC HEMOFILTRATE PEPTIDE HF7665.
CC KAZAL-LIKE 1 (ATYPICAL).
CC KAZAL-LIKE 2.
CC KAZAL-LIKE 3 (ATYPICAL).
CC KAZAL-LIKE 4 (ATYPICAL).
CC KAZAL-LIKE 5 (ATYPICAL).
CC KAZAL-LIKE 6 (ATYPICAL).
CC KAZAL-LIKE 7 (ATYPICAL).
CC KAZAL-LIKE 8 (ATYPICAL).
CC KAZAL-LIKE 9 (ATYPICAL).
CC KAZAL-LIKE 10 (ATYPICAL).
CC KAZAL-LIKE 11 (ATYPICAL).

```

FT DOMAIN 773 828 KAZAL-LIKE 12 (ATYPICAL).
FT DOMAIN 847 903 KAZAL-LIKE 13 (ATYPICAL).
FT DOMAIN 914 970 KAZAL-LIKE 14 (ATYPICAL).
FT DOMAIN 991 1046 KAZAL-LIKE 15.
FT DISULFID 30 66 BY SIMILARITY.
FT DISULFID 44 63 BY SIMILARITY.
FT DISULFID 97 133 BY SIMILARITY.
FT DISULFID 111 130 BY SIMILARITY.
FT DISULFID 119 151 BY SIMILARITY.
FT DISULFID 161 197 BY SIMILARITY.
FT DISULFID 175 194 BY SIMILARITY.
FT DISULFID 225 261 BY SIMILARITY.
FT DISULFID 239 258 BY SIMILARITY.
FT DISULFID 297 333 BY SIMILARITY.
FT DISULFID 311 330 BY SIMILARITY.
FT DISULFID 357 403 BY SIMILARITY.
FT DISULFID 381 400 BY SIMILARITY.
FT DISULFID 437 473 BY SIMILARITY.
FT DISULFID 451 470 BY SIMILARITY.
FT DISULFID 496 532 BY SIMILARITY.
FT DISULFID 510 529 BY SIMILARITY.
FT DISULFID 557 603 BY SIMILARITY.
FT DISULFID 581 600 BY SIMILARITY.
FT DISULFID 632 668 BY SIMILARITY.
FT DISULFID 646 685 BY SIMILARITY.
FT DISULFID 707 743 BY SIMILARITY.
FT DISULFID 721 740 BY SIMILARITY.
FT DISULFID 774 810 BY SIMILARITY.
FT DISULFID 788 807 BY SIMILARITY.
FT DISULFID 849 885 BY SIMILARITY.
FT DISULFID 863 892 BY SIMILARITY.
FT DISULFID 946 949 BY SIMILARITY.
FT DISULFID 930 949 BY SIMILARITY.
FT DISULFID 993 1028 BY SIMILARITY.
FT DISULFID 1006 1025 BY SIMILARITY.
FT DISULFID 1014 1046 BY SIMILARITY.
FT ACT_SITE 46 47 REACTIVE BOND (POTENTIAL).
FT VARIANT 420 420 K->E.
FT 420 420 /FTIG=VAR 015537.
FT CONFLICT 19 27 GCEKXDSLS -> DAASKNEDQ (IN REF. 1).
FT CONFLICT 335 335 A -> V (IN REF. 1).
FT CONFLICT 711 711 R -> Q (IN REF. 1).
SQ SEQUENCE 1064 AA; 120758 MW; 4418063DEA8C341C CRC64;

Query Match 45.7%; Score 37; DB 1; Length 1064;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERTPLTTAAAXAP 14
| | | | |
Db 1048 EESTPTTAAAMP 1060

RESULT 35
ID ICEN XANCT STANDARD; PRT; 1567 AA.
AC F18127;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ice nucleation protein.
GN INAX.
OS Xanthomonas campestris (pv. translucens).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=343;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=X568;
RA Zhao J., Orser C.S.;
RT "Conserved repetition in the ice nucleation gene inax from
Xanthomonas campestris pv. translucens.";

Mol. Gen. Genet. 223:163-166(1990).
CC -1- FUNCTION: Ice nucleation proteins enable bacteria to nucleate
CC crystallization in supercooled water.
CC -1- SUBCELLULAR LOCATION: Outer membrane. (By similarity).
CC -1- DOMAIN: CONTAINS 153 IMPERFECT REPEATS OF THE CONSENSUS
CC OCTAPEPTIDE A-G-Y-G-S-T-L-T; FURTHER ON A 16-RESIDUE AND A
CC REGIONAL 48-RESIDUE PERIODICITY IS SUPERIMPOSED.
CC -1- MISCELLANEOUS: A STRUCTURAL MODEL IS SUGGESTED IN WHICH THE ICE
CC NUCLEATION PROTEIN DISPLAYS A SYMMETRY RELATED TO THAT OF ICE.
CC -1- SIMILARITY: Belongs to the bacterial ice nucleation protein
CC family.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X52970; CAA37140.1; -.
CC DR HSP; P06620; IINA.
CC DR InterPro; IPR000258; Ice_nucleatn.
CC Pfam; PF00818; Ice_nucleation; 81.
CC PRINTS; PR00327; ICENUCLEATN.
CC DR PROSITE; PS00314; ICE_NUCLEATION; 57.
CC SQ SEQUENCE 1567 AA; 152548 MW; C8B451D959ECAD63 CRC64;

Query Match 45.7%; Score 37; DB 1; Length 1567;
Best Local Similarity 58.3%; Pred. No. 2.5e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 ERTPLTTAAAXAP 14
| | | | |
Db 176 EATPVTAAAP 187

RESULT 36
POLG DEN3
ID POLG DEN3 STANDARD; PRT; 3390 AA.
AC P27915;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Genome polypeptide [Contains: Capsid protein C (Core protein); Matrix
DE protein (Envelope glycoprotein M); Major envelope protein E;
DE Nonstructural proteins NS1, NS2, NS4A and NS4B; Protease/helicase
DE (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)
DE (NS5)].
OS Dengue virus type 3.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Flavivirus.
OX NCBI_TaxID=11069;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90266483; PubMed=2345967;
RA Osatomi K., Sumiyoshi H.;
RT "Complete nucleotide sequence of dengue type 3 virus genome RNA.";
Virology 176:643-647(1990).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC precursor polypeptide, commonly with Asp or Glu in the P6
CC position, Cys or Thr in P1 and Ser or Ala in P1'.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -1- SUBUNIT: The virion of this virus is a nucleocapsid covered by a
CC lipoprotein envelope. The envelope consists of two proteins:
CC protein M and glycoprotein E. The nucleocapsid is a complex of
CC protein C and mRNA.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC EMBL; M93130; AAA99437.1; -.
CC PIR; A34774; GNWVD3.
CC MEROPS; S07.002; -.
CC InterPro; IPR009003; Cys Ser_Trypsin.
CC InterPro; IPR001410; DEAD.
CC InterPro; IPR001122; Flavi_capsidC.
CC InterPro; IPR000336; Flavi_glycoprote.
CC InterPro; IPR000069; Flavi_M.
CC InterPro; IPR001157; Flavi_NS1.
CC InterPro; IPR000752; Flavi_NS2A.
CC InterPro; IPR000487; Flavi_NS2B.
CC InterPro; IPR000404; Flavi_NS4A.
CC InterPro; IPR001528; Flavi_NS4B.
CC InterPro; IPR000208; Flavi_NS5.
CC InterPro; IPR002535; Flavi_propep.
CC InterPro; IPR001650; Helicase_C.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR007095; RNA pol DS S7.
CC InterPro; IPR007094; RNA pol_PSVir.
CC InterPro; IPR002877; RnmJ_FtsJ.
CC Pfam; PF01003; Flavi_capsid; 1.
CC Pfam; PF02832; Flavi_Glycop_C; 1.
CC Pfam; PF00869; Flavi_Glycoprot; 1.
CC Pfam; PF00949; Flavi_helicase; 1.
CC Pfam; PF01004; Flavi_M; 1.
CC Pfam; PF00948; Flavi_NS1; 1.
CC Pfam; PF01005; Flavi_NS2A; 1.
CC Pfam; PF01002; Flavi_NS2B; 1.
CC Pfam; PF01350; Flavi_NS4A; 1.
CC Pfam; PF01349; Flavi_NS4B; 1.
CC Pfam; PF00972; Flavi_NS5; 1.
CC Pfam; PF01570; Flavi_propep; 1.
CC Pfam; PF01728; FtsJ_1.
CC Pfam; PF00271; helicase_C; 1.
CC Pfam; PF001556; Flavi_Glycoprote; 1.
CC Pfam; PF001496; Flavi_NS1; 1.
CC SMART; SM00487; DEXDC; 1.
CC SMART; SM00490; HELICC; 1.
CC SMART; SM00690; DEAD_ATP_HELICASE; FALSE NEG.
CC PROSITE; PS00690; DEAD_ATP_HELICASE; RNA-directed RNA polymerase;
KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Hydrolase; Helicase;
KW ATP-binding; Transmembrane; Nonstructural protein.
FT CHAIN 1 114
FT CHAIN 115 205
FT PROPEP 115 205
FT CHAIN 206 280
FT CHAIN 281 773
FT CHAIN 774 1184
FT CHAIN 1185 1343
FT CHAIN 1344 1473
FT CHAIN 1474 2092
FT CHAIN 2093 2378
FT CHAIN 2379 2490
FT CHAIN 2491 3390
FT CHAIN 1667 1674
FT NP_BIND 1667 1674
FT SITE 1758 1761
FT TRANSMEM 46 67
FT TRANSMEM 266 280
FT TRANSMEM 724 746
FT TRANSMEM 753 771
FT TRANSMEM 1156 1175
FT DISULFID 283 310
FT DISULFID 340 396
FT DISULFID 354 385
FT DISULFID 372 401
FT DISULFID 463 563
FT DISULFID 580 611
FT CARBOHYD 183 183
N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```

FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 433 433 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 903 903 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1132 1132 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1188 1188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1661 1661 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2300 2300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2304 2304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2456 2456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2702 2702 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2712 2712 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 3390 AA; 378057 MW; 666B8F70F1E1756E CRC64;
Query Match 45.7%; Score 37; DB 1; Length 3390;
Best Local Similarity 53.8%; Pred. No. 5.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 2 EKKPLTTAAAXAP 14
DB 3315 EDKPTVTWEDVP 3327
|:|:|:|:|
RESULT 37
CAVI_CABEL STANDARD; PRT; 235 AA.
AC Q94051;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Caveolin-1.
GN CAV-1 OR T13F2.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
NCBI_TaxID=6239;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2.
RX MEDLINE=97153022; PubMed=8999956;
RA Tang Z., Okamoto T., Boontrakulpoontawe P., Katada T., Otsuka A.J.,
RA Lisanti M.P.;
RT "Identification, sequence, and expression of an invertebrate caveolin
RT gene family from the nematode Caenorhabditis elegans. Implications
RT for the molecular evolution of mammalian caveolin genes.";
RL J. Biol. Chem. 272:2437-2445(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Swinburne J.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May act as a scaffolding protein within caveolar
CC membranes. Interacts directly with G-protein alpha subunits and
CC can functionally regulate their activity.
CC -!- SUBUNIT: Homooligomer containing 14.15 monomers per oligomer.
CC -!- SUBCELLULAR LOCATION: Membrane protein of caveolae. Potential
CC hairpin-like structure in the membrane (By similarity).
CC -!- SIMILARITY: Belongs to the caveolin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC EMBL; U66405; AAB48388.1; -.
CC EMBL; Z81122; CA503359.1; -.
CC PIR; T24882; T24882.
CC WormPep; T13F2.8; CAV13633.
CC InterPro; IPR001612; Caveolin.
DR
```


CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE016744; AAC03779.1; -
DR HAMAP; MF 00668; -; 1.
DR InterPro; IPR005499; BioW.
DR Pfam; PF03744; BioW; 1.
KW Biotin biosynthesis; Ligase; Magnesium; Complete proteome.
SQ SEQUENCE 228 AA; 25777 MW; 78776158EF87B65B CRC64;

Query Match 44.4%; Score 36; DB 1; Length 228;
Best Local Similarity 46.7%; Pred. No. 51;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 KPTLTAAAPVXN 18
| | | | |
Db 57 KAPIQTLVALPVEN 71

Search completed: October 4, 2004, 15:12:27
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 14:54:19 ; Search time 227 Seconds
(without alignments)
26.409 Million cell updates/sec

Title: US-09-359-426c-2

Perfect score: 81
Sequence: 1 XEEKTPLTAAAXAPVXXN 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	61.7	485	2 Q8GRA4	Q8gra4 porphyromon
2	48	59.3	480	16 Q8ZGS4	Q8zgs4 yersinia pe
3	46	56.8	458	5 Q76672	Q76672 caenorhabdi
4	45	55.6	484	2 Q8GL14	Q8gl14 actinobacil
5	44	54.3	306	16 Q8FKK4	Q8fkk4 escherichia
6	44	54.3	306	16 Q8XGW5	Q8xgw5 salmonella
7	44	54.3	314	2 Q8A119	Q8a119 escherichia
8	44	54.3	427	10 Q23188	Q23188 arabidopsis
9	44	54.3	427	10 Q94A00	Q94a00 arabidopsis
10	43	53.1	483	16 Q82IF1	Q82if1 streptomyce
11	43	53.1	486	2 Q8KKH4	Q8khh4 hyphomicrob
12	43	53.1	506	2 P77924	P77924 pseudomonas
13	43	53.1	1240	12 Q9DWH8	Q9dwh8 rat cytoleg
14	42	51.9	210	2 Q8NU2	Q8nu2 streptococc
15	42	51.9	211	2 Q9X6X8	Q9x6x8 streptococc
16	42	51.9	325	16 Q89EL3	Q89el3 bradyrhizob

17	42	51.9	436	10 Q9FT45	Q9ft45 arabidopsis
18	42	51.9	485	16 Q82TK1	Q82tk1 nitrosomona
19	42	51.9	487	16 Q9RJK9	Q9rjk9 streptomyce
20	42	51.9	751	3 Q86ZL0	Q86zl0 podospora a
21	41	50.6	132	10 Q9SB04	Q9sb04 oryza sativ
22	41	50.6	417	2 Q54272	Q54272 streptomyce
23	41	50.6	516	16 Q8P7W8	Q8p7w8 xanthomonas
24	41	50.6	663	17 Q8ZZX0	Q8zxx0 pyrobaculum
25	41	50.6	920	5 Q8MT63	Q8mt63 drosophila
26	41	50.6	920	5 Q8ML93	Q8ml93 drosophila
27	40	49.4	71	16 Q7U7D9	Q7u7d9 synchococc
28	40	49.4	171	5 Q9U3E3	Q9u3e3 caenorhabdi
29	40	49.4	212	5 Q9U4Y3	Q9u4y3 ceratitis c
30	40	49.4	238	16 Q8PQ46	Q8pq46 xanthomonas
31	40	49.4	238	16 Q8PD69	Q8pd69 xanthomonas
32	40	49.4	252	17 Q9Y984	Q9y984 aeropyrum p
33	40	49.4	258	16 Q8Y5M9	Q8y5m9 listeria mo
34	40	49.4	346	16 Q8ZAJ5	Q8zaj5 streptomyce
35	40	49.4	367	5 Q8IIA1	Q8iia1 drosophila
36	40	49.4	394	16 Q9A4B0	Q9a4b0 caulobacter
37	40	49.4	477	16 Q9RD97	Q9rd97 streptomyce
38	40	49.4	488	16 Q7V5U2	Q7v5u2 prochloroco
39	40	49.4	498	15 Q79665	Q79665 human immun
40	40	49.4	594	5 Q9BLH5	Q9blh5 halocynthia
41	40	49.4	609	11 Q7TWM2	Q7tnw2 mus musculu
42	40	49.4	788	5 Q9W180	Q9w180 drosophila
43	40	49.4	805	5 Q8MLP2	Q8mlp2 drosophila
44	40	49.4	846	13 Q57577	Q57577 cynops pyrr
45	40	49.4	1266	11 Q8CGF1	Q8cgf1 mus musculu

ALIGNMENTS

RESULT 1
Q8GRA4 PRELIMINARY; PRT; 485 AA.
ID Q8GRA4
AC Q8GRA4;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Catalase HP2.
GN KAT.
OS Porphyromonas gulae.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
CX NCBI_TaxID=111105;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFB3492;
RA Nakayama K., Sato K.;
RT "catalase";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR GO; AB083039; BAC20190.1; -
DR GO; GO:0004096; F:catalase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006379; P:response to oxidative stress; IEA.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
SQ SEQUENCE 485 AA; 55610 MW; 6D35ECTA7B59F8AF CRC64;

Query Match 61.7%; Score 50; DB 2; Length 485;
Best Local Similarity 68.8%; Pred. No. 3.4;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 EKTPLTAAAXAPVXXN 18
DB 2 EKNKLTAAAGPVADN 17

OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC	Rhabditidae; Peloderinae; Caenorhabditis.
OX	NCBI_TaxID=6239;
RN	[1]
AC	Q8ZGS4 PRELIMINARY; PRT; 480 AA.
DT	01-MAR-2002 (TRENBLrel. 20, Created)
DT	01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE	Catalase (EC 1.11.1.6).
GN	KATA OR YPO1207 OR KATE OR Y2981.
OS	Yersinia pestis.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Yersinia.
OX	NCBI_TaxID=632;
RN	[1]
RN	SEQUENCE FROM N.A.
RC	STRAIN=CO-92 / Biovar Orientalis;
RC	MEDLINE=21470413; PubMed=11586360;
RA	Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA	Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA	Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Parraga A.M.,
RA	Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA	Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA	Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA	Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT	"Genome sequence of Yersinia pestis, the causative agent of plague.";
RL	Nature 413:523-527(2001).
RN	[2]
RN	SEQUENCE FROM N.A.
RC	STRAIN=KIM5 / Biovar Mediaevalis;
RC	MEDLINE=22137863; PubMed=12142430;
RA	Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA	Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
RA	Perry R.D.;
RT	"Genome sequence of Yersinia pestis KIM.";
RL	J. Bacteriol. 184:4601-4611(2002).
DR	EMBL; AJ414147; CAC90045.1; -
DR	EMBL; AE013900; AAM86532.1; -
DR	PIR; AB0148; AB0148.
DR	HSSP; P42321; IM85.
DR	GO; GO:0004096; F:catalase activity; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0004601; F:peroxidase activity; IEA.
DR	GO; GO:0006118; P:electron transport; IEA.
DR	GO; GO:0006979; P:response to oxidative stress; IEA.
DR	InterPro; IPR002226; Catalase.
DR	Pfam; PF00199; catalase; 1.
DR	PRINTS; PR00067; CATALASE.
DR	ProDom; PD000510; Catalase; 1.
DR	PROSITE; PS00438; CATALASE 2; 1.
KW	Oxidoreductase; Peroxidase; Complete proteome.
SQ	SEQUENCE 480 AA; 55023 NW; 27A72FBAF096347E CRC64;
Query Match	59.3%; Score 48; DB 16; Length 480;
Best Local Similarity	68.8%; Pred. No. 7.5;
Matches	11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY	3 EKTPLTTAAAPVXN 18
Db	4 KKKGTTAGAPVDN 19
RESULT 3	
ID	O76672 PRELIMINARY; PRT; 458 AA.
AC	O76672
DT	01-NOV-1998 (TRENBLrel. 08, Created)
DT	01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE	Hypothetical protein.
GN	H34124.2.
OS	Caenorhabditis elegans.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC	Rhabditidae; Peloderinae; Caenorhabditis.
OX	NCBI_TaxID=6239;
RN	[1]
AC	Q8ZGS4 PRELIMINARY; PRT; 480 AA.
DT	01-MAR-2002 (TRENBLrel. 20, Created)
DT	01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE	Catalase (EC 1.11.1.6).
GN	KATA OR YPO1207 OR KATE OR Y2981.
OS	Yersinia pestis.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC	Enterobacteriaceae; Yersinia.
OX	NCBI_TaxID=632;
RN	[1]
RN	SEQUENCE FROM N.A.
RC	STRAIN=CO-92 / Biovar Orientalis;
RC	MEDLINE=21470413; PubMed=11586360;
RA	Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA	Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA	Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Parraga A.M.,
RA	Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA	Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA	Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA	Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT	"Genome sequence of Yersinia pestis, the causative agent of plague.";
RL	Nature 413:523-527(2001).
RN	[2]
RN	SEQUENCE FROM N.A.
RC	STRAIN=KIM5 / Biovar Mediaevalis;
RC	MEDLINE=22137863; PubMed=12142430;
RA	Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA	Straley S.C., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,
RA	Perry R.D.;
RT	"Genome sequence of Yersinia pestis KIM.";
RL	J. Bacteriol. 184:4601-4611(2002).
DR	EMBL; AJ414147; CAC90045.1; -
DR	EMBL; AE013900; AAM86532.1; -
DR	PIR; AB0148; AB0148.
DR	HSSP; P42321; IM85.
DR	GO; GO:0004096; F:catalase activity; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0004601; F:peroxidase activity; IEA.
DR	GO; GO:0006118; P:electron transport; IEA.
DR	GO; GO:0006979; P:response to oxidative stress; IEA.
DR	InterPro; IPR002226; Catalase.
DR	Pfam; PF00199; catalase; 1.
DR	PRINTS; PR00067; CATALASE.
DR	ProDom; PD000510; Catalase; 1.
DR	PROSITE; PS00438; CATALASE 2; 1.
KW	Oxidoreductase; Peroxidase; Complete proteome.
SQ	SEQUENCE 480 AA; 55023 NW; 27A72FBAF096347E CRC64;
Query Match	59.3%; Score 48; DB 16; Length 480;
Best Local Similarity	68.8%; Pred. No. 7.5;
Matches	11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY	3 EKTPLTTAAAPVXN 18
Db	4 KKKGTTAGAPVDN 19
RESULT 4	
ID	O9RG14 PRELIMINARY; PRT; 484 AA.
AC	O9RG14
DT	01-MAY-2000 (TRENBLrel. 13, Created)
DT	01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT	01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE	Catalase (EC 1.11.1.6).
GN	KATA.
OS	Actinobacillus actinomycetemcomitans (Haemophilus
OS	actinomycetemcomitans)
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC	Pasteurellaceae; Actinobacillus.
OX	NCBI_TaxID=714;
RN	[1]
RN	SEQUENCE FROM N.A.
RC	STRAIN=Y4Na1;
RC	MEDLINE=20042351; PubMed=10572134;
RA	Thomson V.J., Bhattacharjee M.K., Fine D.H., Derbyshire K.M.,
RA	Figurski D.H.;
RT	"Direct Selection of IS903 Transposon Insertions by Use of a Broad-
RT	Host-Range Vector: Isolation of Catalase-Deficient Mutants of
RT	Actinobacillus actinomycetemcomitans.";
RL	J. Bacteriol. 181:7298-7307(1999).
CC	-!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC	SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC	PEROXIDE (BY SIMILARITY).
CC	-!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
DR	EMBL; AF162654; AAF17882.1; -
DR	HSSP; P42321; IM85.
DR	GO; GO:0004096; F:catalase activity; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0004601; F:peroxidase activity; IEA.
DR	GO; GO:0006118; P:electron transport; IEA.

DR GO: GO:0005979; P:response to oxidative stress; IEA.
 DR InterPro; IPR002226; Catalase.
 DR Pfam; PF00199; catalase; 1.
 DR PRINTS; PRO0067; CATALASE.
 DR PROSITE; PS00510; Catalase; 1.
 DR PROSITE; PS00437; CATALASE; 1.
 DR PROSITE; PS00438; CATALASE; 2.
 KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
 SQ SEQUENCE 484 AA; 54961 MW; D0B522AB2557D6CB CRC64;

Query Match 55.6%; Score 45; DB 2; Length 484;
 Best Local Similarity 83.3%; Pred. No. 25;
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 LTTAXAPVXXN 18
 DB 14 LTTAAGAPVDN 25

RESULT 5

Q8FKR4 PRELIMINARY; PRT; 306 AA.
 AC Q8FKR4;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Putative ribokinase.
 GN C0331.
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN:O6.H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
 DR EMBL; AE016756; AAN78819.1; -.
 DR GO: GO:000474; F:ribokinase activity; IEA.
 DR GO: GO:0006014; P:D-ribose metabolism; IEA.
 DR InterPro; IPR002173; PfKB.
 DR InterPro; IPR002139; Ribokinase.
 DR Pfam; PF00294; PfKB; 1.
 DR PRINTS; PRO0990; RIBOKINASE.
 KW Kinase; Complete proteome.
 SQ SEQUENCE 306 AA; 33222 MW; 8A0FBEF68FE493DB CRC64;

Query Match 54.3%; Score 44; DB 16; Length 306;
 Best Local Similarity 47.1%; Pred. No. 23;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 EKTPLTTAAXAPVXXNA 19
 DB 84 EKVPCTSSGVAPIFVNA 100

RESULT 6

Q8XGW5 PRELIMINARY; PRT; 306 AA.
 AC Q8XGW5;
 DT 01-MAR-2002 (TREMELrel. 20, Created)
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Putative carbohydrate kinase (Putative sugar kinase, ribokinase
 family) (EC 2.7.1.15).
 GN T3725 OR STM3793 OR STY3989.
 OS Salmonella typhi, and

OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601, 602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 and Ctr18.";
 RL J. Bacteriol. 185:2330-2337 (2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=L72 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 L72.";
 RL Nature 413:852-856 (2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=CTR18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrall B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 enterica serovar Typhi Ctr18.";
 RL Nature 413:848-852 (2001).
 DR EMBL; AE016846; AA071218.1; -.
 DR EMBL; AE008876; AAL2851.1; -.
 DR GO: GO:000474; F:ribokinase activity; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0006014; P:D-ribose metabolism; IEA.
 DR InterPro; IPR002173; PfKB.
 DR InterPro; IPR002139; Ribokinase.
 DR Pfam; PF00294; PfKB; 1.
 DR PRINTS; PRO0990; RIBOKINASE.
 KW Hypothetical protein; Kinase; Transferase; Complete proteome.
 SQ SEQUENCE 306 AA; 33229 MW; 142A0289E2D1B953 CRC64;

Query Match 54.3%; Score 44; DB 16; Length 306;
 Best Local Similarity 47.1%; Pred. No. 23;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 EKTPLTTAAXAPVXXNA 19
 DB 84 EKVPCTSSGVAPIFVNA 100

RESULT 7

Q9AI19 PRELIMINARY; PRT; 314 AA.
 AC Q9AI19;
 DT 01-JUN-2001 (TREMELrel. 17, Created)
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Hypothetical protein.
 DE Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.

```

OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AL862;
RX MEDLINE=21101044; PubMed=11159989;
RA Lalioui L., Le Bouguenec C.C.;
RT "afa-8 Gene cluster is carried by a pathogenicity island inserted into
RT the tRNA(Phe) of human and bovine pathogenic Escherichia coli
RT isolates."
RL Infect. Immun. 69:937-948(2001).
DR EMBL; AF286671; AAK27331.1; -.
DR HSSP; P05054; 1RK2.
DR GO; GO:0004747; F:ribokinase activity; IEA.
DR GO; GO:0006014; P:D-ribose metabolism; IEA.
DR InterPro; IPR002173; FPKB.
DR InterPro; IPR002139; Ribokinase.
DR Pfam; PF00294; FPKB; 1.
DR PRINTS; PR00990; RIBOKINASE.
KW Hypothetical protein.
SQ
SEQUENCE 314 AA; 34415 MW; 2D5F7342C817E3BB CRC64;

Query Match 54.3%; Score 44; DB 2; Length 314;
Best Local Similarity 47.1%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 EKTLPTTAAAXAPVXNA 19
DB 93 EKVPCTSSGVAPIFVNA 109

RESULT 8
O23188
ID O23188 PRELIMINARY; PRT; 427 AA.
AC O23188
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN C7A10.390 OR AT4G36970.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.; C7A10.390 (GI:4006876).";
RT "Full length cDNA of gene C7A10.390 (GI:4006876).";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY045887; AAK76561.1; -.
DR InterPro; IPR005516; Remorin_C.
DR Pfam; PF03763; Remorin_C; 1.
KW Hypothetical protein.
SQ
SEQUENCE 427 AA; 46798 MW; C8B0192D315DAE41 CRC64;

Query Match 54.3%; Score 44; DB 10; Length 427;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 EKTLPTTAAAXAPV 16
DB 208 EKAEMTTAMQSPV 222

RESULT 10
O821T1
ID O821T1 PRELIMINARY; PRT; 483 AA.
AC O821T1
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative catalase.
GN KAT1L OR SAV3052.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]

```

```

RC SEQUENCE FROM N.A.
RX STRAIN=MA-4690 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AF005033; BAC70763.1; -
DR GO; GO:0004096; F:catalase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Complete proteome.
SQ SEQUENCE 483 AA; 54180 MW; 7353768A63822257 CRC64;

Query Match 53.1%; Score 43; DB 16; Length 483;
Best Local Similarity 69.2%; Pred. No. 54;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 PLTTAAXAPVXN 18
Db 5 PLTTAGAPVDN 17

RESULT 11
ID Q8KKH4 PRELIMINARY; PRT; 486 AA.
AC Q8KKH4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrite reductase precursor.
GN NIRC.
OS Hyphomicrobium denitrificans.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Hyphomicrobiaceae; Hyphomicrobium.
OX NCBI_TaxID=53399;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3151;
RA Kataoka K., Fukui A., Kabayashi M., Yamaguchi K., Suzuki S.;
RT "Cloning and expression of copper-containing nitrite reductase from
RT Hyphomicrobium denitrificans.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB076806; BAC00922.1; -
DR GO; GO:0005507; F:copper ion binding; IEA.
DR GO; GO:0006807; P:nitrogen metabolism; IEA.
DR InterPro; IPR001117; Cu-oxidase.
DR InterPro; IPR001287; CuNO2 reductase.
DR InterPro; IPR008972; Cupredoxin.
DR Pfam; PF00394; Cu-oxidase; 1.
DR PRINTS; PR00655; CUNO2REDTASE.
DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
KW Signal.
FT SIGNAL.
SQ SEQUENCE 1 39
SEQUENCE 486 AA; 52238 MW; 6984E02ECEFF791AF CRC64;

Query Match 53.1%; Score 43; DB 2; Length 486;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 BEKTLTAAAPV 16
Db 53 BEKTDPTAGAPV 67

RESULT 12

```

```

P77924 PRELIMINARY; PRT; 506 AA.
ID P77924;
AC P77924;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Catalase isozyme.
GN KATA
OS Pseudomonas fluorescens.
OG Plasmid pAM10.6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cb36;
RX MEDLINE=21318967; PubMed=11425481;
RA Peters M., Heinaru A., Nurk A.;
RT "Plasmid-encoded catalase KATA, the main catalase of Pseudomonas
RT fluorescens strain Cb36.";
RL FEMS Microbiol. Lett. 200:235-240(2001).
DR EMBL; U72068; BAB17009.1; -
DR HSSP; P42321; IM85.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0004096; F:catalase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Plasmid.
SQ SEQUENCE 506 AA; 57324 MW; F1EAA728C5D41CBE CRC64;

Query Match 53.1%; Score 43; DB 2; Length 506;
Best Local Similarity 71.4%; Pred. No. 56;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TPLTTAAXAPVXN 18
Db 11 TPLTTAGAPVDN 24

RESULT 13
ID Q9DWH8 PRELIMINARY; PRT; 1240 AA.
AC Q9DWH8;
DT 01-VAR-2001 (TrEMBLrel. 16, Created)
DT 01-VAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Pr2.
GN R2.
OS Rat cytomegalovirus (strain Maastricht).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=79700;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Maastricht;
RX MEDLINE=20366325; PubMed=10906222;
RA Vink C., Beuken E., Bruggeman C.A.;
RT "Complete DNA sequence of the rat cytomegalovirus genome.";
RL J. Virol. 74:7656-7665(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Maastricht;
RX MEDLINE=20473137; PubMed=11018281;
RA Gruitjuijzen Y.K., Beuken E., Bruggeman C.A.;
RT "Rat cytomegalovirus R89 is a highly conserved gene which expresses a
RT spliced transcript.";
RL Virus Res. 69:119-130(2000).
RL EMBL; AF232689; AAF99111.1; -

```

SQ SEQUENCE 1240 AA; 125612 MW; 33B6C13DC6A272B0 CRC64;

Query Match 53.1%; Score 43; DB 12; Length 1240;
Best Local Similarity 64.3%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 EEKPLTTAAXAPV 15
||:|||||:
Db 570 EERELTTAGSAPV 583

RESULT 14

Q9RNU2 Q9RNU2 PRELIMINARY; PRT; 210 AA.
AC Q9RNU2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Serum opacity factor precursor (Fragment).
GN S0F4470.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4470-96; TISSUE=Blood;
RA Beall B., Gherardi G., Lovgren M., Tyrrell G., Facklam R., Forwick B.;
RT "Predictions of M serotype, anti-opacity factor type, and highly
RT related strain sets based upon the variable emm and sof gene
RT sequences.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179217; AAD55775.1; -
KW Signal.
FT NON TER 1 1
FT SIGNAL <1 7
FT CHAIN 8 >210
FT NON TER 210 210
SQ SEQUENCE 210 AA; 21220 MW; 199C99EFCF260B6A CRC64;

Query Match 51.9%; Score 42; DB 2; Length 210;
Best Local Similarity 57.1%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 EEKPLTTAAXAPV 15
||:|||||:
Db 103 EEKPKVTSPTV 116

RESULT 15

Q9X6X8 Q9X6X8 PRELIMINARY; PRT; 211 AA.
AC Q9X6X8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Serum opacity factor precursor (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SS1401 type PT1658;
RA Beall B., Gherardi G.;
RT "The relation of Streptococcus pyogenes sof and emm gene sequence
RT types to genetically distinct strain sets.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF154330; AAD36988.1; -
KW Signal.
FT NON TER 1 1
FT SIGNAL <1 8
FT CHAIN 9 >211
FT SERUM OPACITY FACTOR.

FT NON TER 211 211
SQ SEQUENCE 211 AA; 21367 MW; F5475DDC6A084FE6 CRC64;

Query Match 51.9%; Score 42; DB 2; Length 211;
Best Local Similarity 61.5%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 EEKPLTTAAXAP 14
||:|||||:
Db 98 EEKPKTATSSAP 110

RESULT 16

Q89EL3 Q89EL3 PRELIMINARY; PRT; 325 AA.
AC Q89EL3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE BL7060 protein.
GN BL7060.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpou S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005960; BAC52325.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6.
DR InterPro; IPR001092; HLF_basic.
DR Pfam; PF00892; DUF6; 2.
DR PROSITE; PS00038; HLH_1; 1.
KW Complete Proteome.
SQ SEQUENCE 325 AA; 35291 MW; 42AB546310432A94 CRC64;

Query Match 51.9%; Score 42; DB 16; Length 325;
Best Local Similarity 57.1%; Pred. No. 53;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 TELTTAAXAPVXN 18
||:|||||:
Db 13 TFWTTASPAPPASN 26

RESULT 17

Q8FT45 Q8FT45 PRELIMINARY; PRT; 436 AA.
AC Q8FT45;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (GPI-anchored protein).
GN T25B15.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Alcaraz J.P., Clabault G., Cortet A., Mache R., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

```

RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RC STRAIN=cv. Columbia;
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Arabidopsis thaliana full-length cDNA.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132972; CAC07928.1; -;
DR EMBL; AK117608; BAC42264.1; -;
DR PIR; T46107; T46107.
DR GO; GO:0004379; F:glycylpeptide N-tetradecanoyltransferase ac. . .; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006499; P:N-terminal protein myristoylation; IEA.
DR InterPro; IPR000782; BIGH3_FAS1.
DR InterPro; IPR000903; Nmt.
DR Pfam; PF02469; Fasciclin; 1.
DR SMART; SM00554; FAS1; 2.
DR PROSITE; PS02013; FAS1; 2.
DR PROSITE; PS00976; NMT_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 436 AA; 48073 MW; 83FD94879F9855DF CRC64;

Query Match 51.9%; Score 42; DB 10; Length 436;
Best Local Similarity 50.0%; Pred. No. 72;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 EKTPLTTAAAPVXXN 19
DB 395 EKTPEVKKTGVVXXA 412.

RESULT 18
Q82TK1 PRELIMINARY; PRT; 485 AA.
ID Q82TK1
AC Q82TK1
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidaigo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Mark L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
DR EMBL; AL93105; CAB58320.1; -.
DR HSSP; P42321; 1M85.
DR GO; GO:0004096; F:catalase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR PRODOM; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Complete proteome.
SQ SEQUENCE 485 AA; 55482 MW; 875EA3A6D751A6F3 CRC64;

```

```

Query Match 51.9%; Score 42; DB 16; Length 485;
Best Local Similarity 52.9%; Pred. No. 80;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 EKTPLTTAAAPVXXN 18
DB 3 DAKKLTTSAGAPVADN 19

RESULT 19
Q9RJK9 PRELIMINARY; PRT; 487 AA.
ID Q9RJK9
AC Q9RJK9
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Catalase (EC 1.11.1.6).
DE KATA OR SC00379 OR SCF62.05.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OC NCBI_taxid=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidaigo J., Hornsby T., Howarth S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Oliver K., O'Neill S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
DR EMBL; AL93105; CAB58320.1; -.
DR HSSP; P42321; 1M85.
DR GO; GO:0004096; F:catalase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004601; F:peroxidase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006979; P:response to oxidative stress; IEA.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR PRODOM; PD000510; Catalase; 1.

```

DR PROSITE; PS00437; CATALASE_1; 1.
 DR PROSITE; PS00438; CATALASE_2; 1.
 KW Heme; Hydrogen peroxidase; Iron; Oxidoreductase; Peroxidase;
 KW Complete proteome.
 SQ SEQUENCE 487 AA; 55116 MW; 9D3334889EAF60B7 CRC64;
 Query Match 51.9%; Score 42; DB 16; Length 487;
 Best Local Similarity 69.2%; Pred. No. 80;
 Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 6 PLTTAAAPVXN 18
 |||||
 DB 8 PLTTAGAPVDN 20
 |||||
 RESULT 20
 Q86ZLO PRELIMINARY; PRT; 751 AA.
 AC Q86ZLO;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Similar to conserved oligomeric golgi complex component 6 of Mus musculus.
 DE Musculus.
 OS Podospira anserina.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.
 OX NCBI_TaxID=5145;
 [1]
 RN SEQUENCE FROM N.A.
 RA Genoscope;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BX088700; CAD60710.1; -.
 SQ SEQUENCE 751 AA; 83366 MW; B27A29A65A863341 CRC64;
 Query Match 51.9%; Score 42; DB 3; Length 751;
 Best Local Similarity 52.9%; Pred. No. 1.2e+02;
 Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 3 EKTPLTTAAAPVXNA 19
 |||||
 DB 649 EKVELLEAAQALTOA 665
 |||||
 RESULT 21
 Q9SE04 PRELIMINARY; PRT; 132 AA.
 AC Q9SE04;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Copper chaperone homolog CCH.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eriactoidae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 [1]
 RN SEQUENCE FROM N.A.
 RA Himelblau E., Mira H., Lin S.J., Culotta V.C., Penarrubia L.,
 RA Masino R.M.;
 RT Identification of a functional homolog of the yeast copper
 RT homeostasis gene ATR1 from Arabidopsis.
 RL Plant Physiol. 117:1227-1234 (1998).
 [2]
 RN SEQUENCE FROM N.A.
 RA Mira H., Penarrubia L.;
 RT "Copper chaperone from Oryza sativa."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF198626; AAF15285.1; -.
 DR PIR; T50779; T50779.
 DR HSSP; O00244; 1FE0.

DR Gramene; Q9SE04; -.
 DR GO; GO:0046872; P:metal ion binding; IEA.
 DR GO; GO:0030001; P:metal ion transport; IEA.
 DR InterPro; IPR006121; HeavyMe_transpt.
 DR InterPro; IPR006191; Metal_bind.
 DR Pfam; PF00403; HMA; 1.
 DR PROSITE; PS00846; HMA_2; 1.
 SQ SEQUENCE 132 AA; 13094 MW; 7176EF95350A8231 CRC64;
 Query Match 50.6%; Score 41; DB 10; Length 132;
 Best Local Similarity 50.0%; Pred. No. 32;
 Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
 QY 2 EKTPLTTAAAPVXNA 19
 |||||
 DB 97 EAAPPTTAAEAPAA 114
 |||||
 RESULT 22
 Q54272 PRELIMINARY; PRT; 417 AA.
 AC Q54272;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 OS Streptomyces hygroscopicus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1912;
 [1]
 RN SEQUENCE FROM N.A.
 RA STRAIN-SF1293;
 RX MEDLINE=95309717; PubMed=7789803;
 RA Hidaka T., Hidaka M., Kuzuyama T., Seto H.;
 RT "Sequence of a p-methyltransferase encoding gene isolated from a
 RT biolaphos-producing streptomyces hygroscopicus."
 RL Gene 158:149-150 (1995).
 CC -!- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
 DR EMBL; D37877; BAA07116.1; -.
 DR PIR; PC4031; PC4031.
 DR HSSP; P56533; 1A4S.
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyde_dehydr.
 DR Pfam; PF00171; aldehyd; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 KW Hypothetical protein; Oxidoreductase.
 SQ SEQUENCE 417 AA; 43406 MW; 9B8714F4FB654B47 CRC64;
 Query Match 50.6%; Score 41; DB 2; Length 417;
 Best Local Similarity 52.9%; Pred. No. 1e+02;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
 QY 3 EKTPLTTAAAPVXNA 19
 |||||
 DB 117 EKTPLTALAPAELEIA 133
 |||||
 RESULT 23
 Q8P7W8 PRELIMINARY; PRT; 516 AA.
 AC Q8P7W8;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE PILY1 protein.
 GN PILY1 OR XCC2489.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=3340;
 [1]
 RN


```

RC SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=20202145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Canarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Forighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Kasuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;; of two Xanthomonas pathogens with differing
RT host specificities;
RL Nature 417:459-463(2002).
DR EMBL; AE012360; AA041763.1; -.
KW Complete proteome.
SQ SEQUENCE 516 AA; 53416 MW; 88A17A8B40B36578 CRC64;

Query Match 50.6%; Score 41; DB 16; Length 516;
Best Local Similarity 72.7%; Pred. No. 1.3e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 TPLTTAXAPV 15
DB 454 TPLTSGNAPV 464

ID Q82ZX0 PRELIMINARY; PRT; 663 AA.
AC Q82ZX0;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Conserved within P. aerophilum.
GN PAE0040.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009747; AAL62519.1; -.
DR InterPro; IPR001064; CysTallin.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
KW Complete proteome.
SQ SEQUENCE 663 AA; 71372 MW; 216F07B921B78B76 CRC64;

Query Match 50.6%; Score 41; DB 17; Length 663;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 9; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 EKTPLTTAXAPVVXIA 19
DB 617 ESATPATTATPKEDNA 634

RESULT 25
Q8MT63

```

```

ID Q8MT63 PRELIMINARY; PRT; 920 AA.
AC Q8MT63;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE RE03056P.
GN CG30480 OR CG18373.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Cealniker S.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV118357; AM48386.1; -.
DR FlyBase; FBgn0050480; CG30480.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002052; N6 Mbase.
DR PROSITE; PS00092; N6 MTASE; 1.
SQ SEQUENCE 920 AA; 101042 MW; 34FF1CCBD293AF79 CRC64;

Query Match 50.6%; Score 41; DB 5; Length 920;
Best Local Similarity 53.3%; Pred. No. 2.3e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKTPLTTAXAPVV 16
DB 695 EVTPYNTAPHSPIV 709

RESULT 26
Q8ML93
ID Q8ML93 PRELIMINARY; PRT; 920 AA.
AC Q8ML93;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE CG30480-PA.
GN CG30480 OR CG18373.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Cealniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoche C., Baldwin D.,
RA Ballif R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

```

RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Klamm B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobaraj C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celisner S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Bantz J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champs M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaninger J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Carlson J.W., Celisner S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E., S.E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.B.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003813; AAME8547.1; -;
DR FlyBase; FB00050480; CG30480.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0008170; F:N-methyltransferase activity; IEA.
DR GO; GO:0006306; P:DNA methylation; IEA.
DR InterPro; IPR002052; N6 Mtaase.
DR PROSITE; PS00092; N6 MTAASE; 1.
SQ SEQUENCE 920 AA; 101070 MW; 4892AC24257C5203 CRC64;

Query Match 50.6%; Score 41; DB 5; Length 920;
Best Local Similarity 53.3%; Pred. No. 2.3e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 EEKTLTAAAXAPVV 16

DB 695 BEVTPYNTAPHSPIV 709
RESULT 27
OYU7D9
ID O7U7D9 PRELIMINARY; PRT; 71 AA.
AC O7U7D9;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical.
GN SYNW1046.
OS *Synechococcus* sp. (strain WH8102).
OX Bacteria; Cyanobacteria; Chroococcales; *Synechococcus*.
OX NCBI_TaxID=84588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825697; PubMed=12917641;
RA Palenik B., Brahmsha B., Larimer F.W., Land M., Hauser L., Chain P.,
RA Lamerdin J., Regala W., Allen E.E., McCarren J., Paulsen I.,
RA Dufresne A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine *Synechococcus*.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569691; CAE07561.1; -;
SQ SEQUENCE 71 AA; 7628 MW; AF8EDF9EFBB2D28E CRC64;
Query Match 49.4%; Score 40; DB 16; Length 71;
Best Local Similarity 88.9%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 6 PLTTAAXAP 14
DB 61 PLTTAASAP 69
RESULT 28
OYU3E3
ID O9U3E3 PRELIMINARY; PRT; 171 AA.
AC O9U3E3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE F56F12.1 protein.
GN F56F12.1.
OS *Caenorhabditis elegans*.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; *Caenorhabditis*.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Ainscough R.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode *C. elegans*: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z82273; CAB54978.1; -;
DR PIR; T31478; T31478.
DR WormPep; F56F12.1; CE23785.
DR InterPro; IPR003127; Sorb.
DR SMART; SM00459; Sorb; 1.
SQ SEQUENCE 171 AA; 18844 MW; D782482A786F7941 CRC64;
Query Match 49.4%; Score 40; DB 5; Length 171;
Best Local Similarity 53.3%; Pred. No. 61;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
OY 5 TPLTTAAXAPVVXNA 19

DR	InterPro: IPR002942; S4.	DR	InterPro: IPR002942; S4.
DR	PFam: PF01479; S4; 1.	DR	PFam: PF01479; S4; 1.
DR	SMART: SM00363; S4; 1.	DR	SMART: SM00363; S4; 1.
DR	PROSITE: PS00889; S4; 1.	DR	PROSITE: PS00889; S4; 1.
KW	Hypothetical protein; Complete proteome.	KW	Hypothetical protein; Complete proteome.
SQ	SEQUENCE 258 AA; 29682 MW; AB62C55270CFD90F CRC64;	SQ	SEQUENCE 258 AA; 29682 MW; AB62C55270CFD90F CRC64;
Query Match	49.4%; Score 40; DB 16; Length 258;	Query Match	49.4%; Score 40; DB 16; Length 258;
Best Local Similarity	57.1%; Pred. No. 93;	Best Local Similarity	57.1%; Pred. No. 93;
Matches	8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;	Matches	8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY	3 EKTPLTTAAAXAPVV 16	QY	3 EKTPLTTAAAXAPVV 16
DB	158 EETELTNAVYAPVI 171	DB	158 EETELTNAVYAPVI 171
RESULT 34		RESULT 35	
Q82AJ5	PRELIMINARY; PRT; 346 AA.	Q811A1	PRELIMINARY; PRT; 367 AA.
AC	Q82AJ5	AC	Q811A1
DT	01-JUN-2003 (TrEMBLrel. 24, Created)	DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)	DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Putative secreted protein.	DE	Eve-PA.
GN	SAV6062.	GN	EVE.
OS	Streptomyces avermitilis.	OS	Streptomyces avermitilis.
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC	Streptomycineae; Streptomycetaceae; Streptomyces.	OC	Streptomycineae; Streptomycetaceae; Streptomyces.
OX	NCBI_TaxID=33903;	OX	NCBI_TaxID=33903;
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;	RC	STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX	MEDLINE=21477403; PubMed=11572948;	RX	MEDLINE=21477403; PubMed=11572948;
RA	Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,	RA	Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA	Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,	RA	Sakaki Y., Hattori M., Omura S.;
RA	Kikuchi H., Shiba T., Sakaki Y., Hattori M.;	RA	"Complete genome sequence and comparative analysis of the industrial
RT	"Genome sequence of an industrial microorganism Streptomyces	RT	microorganism Streptomyces avermitilis."
RT	avermitilis: deducing the ability of producing secondary	RT	metabolites."
RT	metabolites."	RT	metabolites."
RL	Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).	RL	Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN	[2]	RN	[2]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RC	STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;	RC	STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX	MEDLINE=22608306; PubMed=12692562;	RX	MEDLINE=22608306; PubMed=12692562;
RA	Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,	RA	Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA	Sakaki Y., Hattori M., Omura S.;	RA	Sakaki Y., Hattori M., Omura S.;
RT	"Complete genome sequence and comparative analysis of the industrial	RT	microorganism Streptomyces avermitilis."
RT	microorganism Streptomyces avermitilis."	RT	microorganism Streptomyces avermitilis."
RL	Nat. Biotechnol. 21:526-531(2003).	RL	Nat. Biotechnol. 21:526-531(2003).
DR	EMBL: AP005045; BAC73773.1;	DR	EMBL: AP005045; BAC73773.1;
DR	InterPro: IPR000664; Peptidase_C40.	DR	InterPro: IPR000664; Peptidase_C40.
DR	PFam: PF00877; NLP_C60; 1.	DR	PFam: PF00877; NLP_C60; 1.
KW	Complete proteome.	KW	Complete proteome.
SQ	SEQUENCE 346 AA; 37006 MW; 1717D483B9E283FD CRC64;	SQ	SEQUENCE 346 AA; 37006 MW; 1717D483B9E283FD CRC64;
Query Match	49.4%; Score 40; DB 16; Length 346;	Query Match	49.4%; Score 40; DB 16; Length 346;
Best Local Similarity	66.7%; Pred. No. 1.2e+02;	Best Local Similarity	66.7%; Pred. No. 1.2e+02;
Matches	10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;	Matches	10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY	5 TPLTTAAAXAPVVXNA 19	QY	5 TPLTTAAAXAPVVXNA 19
DB	16 TVLTTAAAAAVVLSA 30	DB	16 TVLTTAAAAAVVLSA 30
RESULT 35		RESULT 35	
Q811A1	PRELIMINARY; PRT; 367 AA.	Q811A1	PRELIMINARY; PRT; 367 AA.
ID	Q811A1	ID	Q811A1
AC	Q811A1	AC	Q811A1
DT	01-MAR-2003 (TrEMBLrel. 23, Created)	DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)	DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Eve-PA.	DE	Eve-PA.
GN	EVE.	GN	EVE.

DR Pfam; PF03929; DUF337; 1.
SQ Hypochemical protein; Complete proteome.
KW SEQUENCE 394 AA; 42675 MW; 8B11178E174D18D CRC64;

Query Match 49.4%; Score 40; DB 16; Length 394;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps

OY 4 KTPLTTAAAXAPVVXNA 19
||||| :||
Db 315 KTPLTTLTTPALIDA 330

RESULT 37
Q9RD97 PRELIMINARY; PRT; 477 AA.
ID AC Q9RD97
AD Q9RD97;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative transmembrane efflux protein.
DN SCO2854 OR SCE20.28.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID:1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RL Seeger K.J., Harris D.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RL Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=97000351; PubMed=8843436;
RX Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RA 'A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.';
RT Mol. Microbiol. 21:77-96 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21896410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinovitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RT Nature 417:141-147 (2002).
DR EMBL; AL939114; CAB65584.1; --
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005829; Sug. transporter.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Transmembrane; Complete proteome.
SQ SEQUENCE 477 AA; 47686 MW; 20797B126B148090 CRC64;

Query Match 49.4%; Score 40; DB 16; Length 477;

```

Best Local Similarity 80.0%; Pred. No. 1.7e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 8; Conservative 0;

Qy 5 TPJTAAAXAP 14
Db 19 TPJTAAAXAP 28

RESULT 38
Q7V5U2 PRELIMINARY; PRT; 488 AA.
AC Q7V5U2;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DE ATP synthase beta subunit, central region:ATP synth. . . (EC 3.6.3.14)
GN ATPB OR PMT1451.
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Algren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572099, CAE21626.1; -.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 488 AA; 52136 MW; 0931106ABDE5668D CRC64;

Query Match 49.4%; Score 40; DB 16; Length 488;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 EEXTPLTTAAAXAPV 15
Db 111 DEQGPVTTDATAPI 124

RESULT 39
Q79665 PRELIMINARY; PRT; 498 AA.
AC Q79665;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Gag polyprotein.
GN GAG.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94149848; PubMed=8107219;
RA Gurtler L.G., Hauser P.H., Eberle J., von Brunn A., Knapp S.,
RA Zekeng L., Isagau J.M., Kaptue L.;
RT "A new subtype of human immunodeficiency virus type 1 (MVP-5180) from Cameroon.";
RL J. Virol. 68:1581-1585(1994).
DR EMBL; L20571; AAA44859.1; -.
DR HSSP; P05888; IAAF.
DR GO; GO:0019012; C:viron; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016032; P:viral life cycle; IEA.
DR InterPro; IPR000721; Gag_p24.

```

```

DR InterPro; IPR000071; Retrovir_p17.
DR InterPro; IPR008916; Retrov_capsid_C.
DR InterPro; IPR008919; Retrov_capsid_N.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF00540; Gag_p17; 1.
DR Pfam; PF00607; Gag_p24; 1.
DR Pfam; PF00098; zf-CCHC; 2.
DR PRINTS; PR00939; C2HCZNFINGER.
DR PRINTS; PR00234; HIVMATRIX.
DR SMART; SM00343; Znf_C2HC; 2.
DR PROSITE; PSS0158; ZF_CCHC; 2.
KW AIDS; Core protein; Polyprotein.
SQ SEQUENCE 498 AA; 55498 MW; 6DE3A29B4E8A17D4 CRC64;

Query Match 49.4%; Score 40; DB 15; Length 498;
Best Local Similarity 44.4%; Pred. No. 1.8e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 EEXTPLTTAAAXAPVXNA 19
Db 119 EETSPTSQNYPIVNA 136

RESULT 40
Q9BLH5 PRELIMINARY; PRT; 594 AA.
AC Q9BLH5;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE HRETR-1 protein.
GN HRETR-1.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyuridae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RA Yagi K., Makabe K.W.;
RT "New neural marker genes expressing in subsets of embryonic neural cells of the ascidian, Halocynthia roretzi.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047035; BAB40781.1; -.
DR HSSP; P19339; 2SKL.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR002343; Hud_Sxl_RNA.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 3.
DR PRINTS; PR00561; HUDSLRNA.
DR SMART; SM00360; RRM; 3.
DR PROSITE; PSS0102; RRM; 3.
SQ SEQUENCE 594 AA; 63686 MW; 193CDE2B14216920 CRC64;

Query Match 49.4%; Score 40; DB 5; Length 594;
Best Local Similarity 46.7%; Pred. No. 2.2e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 5 TPJTAAAXAPVXNA 19
Db 483 TPJTAAAXAPVXNA 497

Search completed: October 4, 2004, 15:16:33
Job time : 235 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 14:15:35 ; Search time 230 Seconds
(without alignments)
23.341 Million cell updates/sec

Title: US-09-359-426C-2

Perfect score: 81

Sequence: 1 XEEXTPLTAAAXAPVWXA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	76	93.8	19	2	AAW64480	1	AAW64480 P. aerugi
2	76	93.8	19	4	ABE6062	2	ABE6062 Pseudomon
3	53	65.4	495	6	ABU40750	3	ABU40750 Protein e
4	48	59.3	480	6	ABU49849	4	ABU49849 Protein e
5	44	54.3	291	3	AAAG30091	5	AAAG30091 Arabidops
6	44	54.3	292	3	AAAG30090	6	AAAG30090 Arabidops
7	44	54.3	306	4	ABE52865	7	ABE52865 Escherich
8	44	54.3	306	4	ABE67588	8	ABE67588 Amino aci
9	44	54.3	427	3	AAAG30089	9	AAAG30089 Arabidops
10	44	54.3	584	6	ABU19765	10	ABU19765 Protein e
11	42	51.9	172	3	AAAG3604	11	AAAG3604 Arabidops
12	42	51.9	215	3	AAAG3603	12	AAAG3603 Arabidops
13	42	51.9	332	3	AAAG3602	13	AAAG3602 Arabidops
14	41	50.6	132	6	ABU58158	14	ABU58158 Rice stre
15	41	50.6	517	6	ABU23415	15	ABU23415 Protein e
16	41	50.6	1012	4	ABE65926	16	ABE65926 Drosophil
17	40	49.4	258	5	ABE48482	17	ABE48482 Listeria
18	40	49.4	445	4	AAU36108	18	AAU36108 Klebsiell
19	40	49.4	498	2	AAAR51691	19	AAAR51691 HIV-type
20	40	49.4	498	2	AAW93076	20	AAW93076 HIV isola
21	40	49.4	500	6	ABJ37450	21	ABJ37450 Benzodiaz
22	40	49.4	788	4	ABE60583	22	ABE60583 Drosophil
23	40	49.4	788	7	AAE38171	23	AAE38171 Fruit fly
24	40	49.4	1473	4	ABE68024	24	ABE68024 Drosophil
25	39.5	48.8	1793	4	ABE59613	25	ABE59613 Drosophil

ALIGNMENTS

RESULT 1

AAW64480

ID AAW64480 standard; peptide; 19 AA.

XX AAW64480;

DT 20-OCT-1998 (first entry)

DE P. aeruginosa protein antigen Pa60 N-terminal peptide fragment.

KW Antigen; Pa60; diagnosis; detection; cystic fibrosis; vaccine; immunogen;

RW infection; treatment.

XX Pseudomonas aeruginosa.

PH Key Location/Qualifiers

FT Misc-difference 1 /label= unknown

FT Misc-difference 12 /label= unknown

FT Misc-difference 17 /label= unknown

XX WO9832769-A1.

XX 30-JUL-1998.

XX 26-JAN-1998; 98WO-GB000217.

XX 24-JAN-1997; 97GB-00001489.

PA (AUSP-) AUSPHARM INT LTD.

PA (CHAP/) CHAPMAN P W.

XX Cripps AW, Kyd J, Dunkley M, Ciancy RL;

WPI; 1998-427879/36.

XX Protein antigen from Pseudomonas aeruginosa and its antigenic fragments -
useful diagnostically to detect specific antibodies, particularly in
patients with cystic fibrosis, and as vaccines.

PS Claim 3; Page 2; 23pp; English.

CC This peptide is the N-terminal fragment of a novel Pseudomonas aeruginosa
protein antigen, Pa60. This fragment could be used for diagnostic
detection of P. aeruginosa by forming complexes with specific antibodies,
particularly in patients with cystic fibrosis (by analysis of mucus, e.g.

Ag939690 Arabidops
Abg26339 Novel hum
Ada35852 Acinetoba
Adb08554 Alloococ
Abg77349 Selected
Abj11251 Yeast sel
Ag91909 C glutami
Aae22281 Murine SP
Aae22282 Murine no
Aae22283 Murine tu
Abu38941 Protein e
Ade31679 Human 707
Abb67966 Drosophil
Adc32844 Human nov
Abb1511 Drosophil
Aab94775 Human pro
Abb1876 Double st
Aau81227 Human lun
Aag91455 C glutami
Abp73930 Candida a

26 39 48.1 128 3 AAG39690
27 39 48.1 137 4 ABG26339
28 39 48.1 145 6 ADA35852
29 39 48.1 154 6 ADB08554
30 39 48.1 159 5 ABG77349
31 39 48.1 159 5 ABJ11251
32 39 48.1 233 4 AAG91909
33 39 48.1 321 5 AAE22281
34 39 48.1 395 5 AAE22283
35 39 48.1 395 5 AAE22282
36 39 48.1 484 6 ABU38941
37 39 48.1 2080 7 ADE31679
38 38.5 47.5 842 4 ABB67966
39 38 46.9 203 7 ADC32844
40 38 46.9 225 4 ABB1511
41 38 46.9 229 4 AAB94775
42 38 46.9 229 5 ABB1876
43 38 46.9 229 5 AAU81227
44 38 46.9 261 4 AAG91455
45 38 46.9 276 5 ABP73930

CC in saliva), or in vaccines or immunogenic compositions to treat or
 CC prevent infection by *P. aeruginosa*
 XX
 SQ Sequence 19 AA;

Query Match 93.8%; Score 76; DB 2; Length 19;
 Best Local Similarity 100.0%; Pred. No. 5.1e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 BEKPLTTAAAXAPVVXNA 19
 |||||
 Db 2 BEKPLTTAAAXAPVVXNA 19
 |||||

RESULT 2
 AAB69062
 ID AAB69062 standard; peptide; 19 AA.
 AC AAB69062;
 XX
 DT 18-APR-2001 (first entry)
 DE Pseudomonas aeruginosa protein N-terminal peptide.
 KW Pseudomonas aeruginosa; chitinase; groEL; chiA; antigen; vaccine;
 KW diagnosis; detection; infection; immune response.
 XX
 CS Pseudomonas aeruginosa.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "unspecified"
 FT Misc-difference 12 /note= "unspecified"
 FT Misc-difference 17 /note= "unspecified"
 FT Misc-difference 17 /note= "unspecified"
 XX
 FN WO200102577-A1.
 XX
 PD 11-JAN-2001.
 XX
 PF 03-JUL-2000; 2000WO-GB002554.
 XX
 PR 01-JUL-1999; 99GB-00015419.
 XX
 PA (PROV-) PROVALIS UK LTD.
 XX
 PI Smith CJ, Thompson SE, Smith MW, Peek K, Sizer PJH, Wilkinson MC;
 XX
 DR WPI; 2001-080988/09.
 XX
 PT Antigenic Pseudomonas aeruginosa proteins, useful in the detection and/or
 PT diagnosis of *P. aeruginosa* infections and for producing vaccines against
 PT *P. aeruginosa*.
 XX
 PS Disclosure; Page 2; 129pp; English.
 XX

The present invention describes antigenic Pseudomonas aeruginosa proteins
 (Pi). The *P. aeruginosa* proteins have antibacterial activity and can be
 used in vaccines and as antagonists. The proteins or their fragments, or
 antibodies are useful in the detection and/or diagnosis of *P. aeruginosa*.
 They are also useful for producing a vaccine and inducing an immune
 response against *P. aeruginosa* infection. An agent capable of
 antagonising, inhibiting or otherwise interfering with the function or
 expression of Pi are useful in the manufacture of a medicament for the
 treatment or prophylaxis of *P. aeruginosa* infections. The present
 sequence represents a probable *P. aeruginosa* protein N-terminal peptide
 sequence from the present invention

Sequence 19 AA;
 Query Match 93.8%; Score 76; DB 4; Length 19;
 Best Local Similarity 100.0%; Pred. No. 5.1e-06;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 BEKPLTTAAAXAPVVXNA 19
 |||||
 Db 2 BEKPLTTAAAXAPVVXNA 19
 |||||
 RESULT 3
 ABU40750
 ID ABU40750 standard; protein; 495 AA.
 XX
 AC ABU40750;
 XX
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #26277.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Proteus sp.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA44620.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 68674; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC

CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC parent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 495 AA;
Query Match 65.4%; Score 53; DB 6; Length 495;
Best Local Similarity 75.0%; Pred. No. 1.6;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 3 EKTPLTTAAAXAPVYN 18
Db 13 EKKLLTTAAGAPVVDN 28
RESULT 4
ABU49849
ID ABU49849 standard; protein; 480 AA.
XX
AC ABU49849;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #35376.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Versinia pestis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002;
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA53719.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids, required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 77773; 1765pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 480 AA;
Query Match 59.3%; Score 48; DB 6; Length 480;
Best Local Similarity 68.8%; Pred. No. 11;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 3 EKTPLTTAAAXAPVYN 18
Db 4 KKKGLTTAAGAPVVDN 19
RESULT 5
AAG30091
ID AAG30091 standard; protein; 291 AA.
XX
AC AAG30091;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35912.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130911P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136382P.
PR 28-MAY-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137502P.
PR 08-JUN-1999; 99US-0137724P.
PR 10-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 16-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 28-JUN-1999; 99US-0140655P.
PR 29-JUN-1999; 99US-0140823P.
PR 30-JUN-1999; 99US-0140991P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 06-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142300P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 18-JUL-1999; 99US-0144086P.
PR 18-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145919P.
PR 02-AUG-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158039P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.

```

PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 26-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 54.3%; Score 44; DB 3; Length 291;
Best Local Similarity 60.0%; Pred.No. 30;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 2 BEKTLTAAAPV 16
Db 72 BEKAETTMQSPV 86
|||||:|||||

RESULT 6
AAG30090
ID AAG30090 standard; protein; 292 AA.
AC AAG30090;
XX
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 35911.
XX
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX
OS Arabidopsis thaliana.
XX
XX
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134222P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140911P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 18-JUL-1999; 99US-0144086P.
PR 18-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.

```

```

PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147284P.
PR 05-AUG-1999; 99US-0147302P.
PR 06-AUG-1999; 99US-0147192P.
PR 07-AUG-1999; 99US-0147260P.
PR 08-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147416P.
PR 10-AUG-1999; 99US-0147493P.
PR 11-AUG-1999; 99US-0148171P.
PR 12-AUG-1999; 99US-0148319P.
PR 13-AUG-1999; 99US-0148341P.
PR 14-AUG-1999; 99US-0148565P.
PR 15-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 19-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 21-AUG-1999; 99US-0149902P.
PR 22-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 24-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150586P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 28-AUG-1999; 99US-0151066P.
PR 29-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 02-SEP-1999; 99US-0152363P.
PR 03-SEP-1999; 99US-0153070P.
PR 04-SEP-1999; 99US-0153758P.
PR 05-SEP-1999; 99US-0154018P.
PR 06-SEP-1999; 99US-0154039P.
PR 07-SEP-1999; 99US-0154779P.
PR 08-SEP-1999; 99US-0155139P.
PR 09-SEP-1999; 99US-0155486P.
PR 10-SEP-1999; 99US-0155659P.
PR 11-SEP-1999; 99US-0156458P.
PR 12-SEP-1999; 99US-0156596P.
PR 13-SEP-1999; 99US-0157117P.
PR 14-SEP-1999; 99US-0157753P.
PR 15-SEP-1999; 99US-0157855P.
PR 16-SEP-1999; 99US-0158029P.
PR 17-SEP-1999; 99US-0158232P.
PR 18-SEP-1999; 99US-0158369P.
PR 19-SEP-1999; 99US-0159293P.
PR 20-SEP-1999; 99US-0159294P.
PR 21-SEP-1999; 99US-0159295P.
PR 22-SEP-1999; 99US-0159330P.
PR 23-SEP-1999; 99US-0159331P.
PR 24-SEP-1999; 99US-0159637P.
PR 25-SEP-1999; 99US-0159638P.
PR 26-SEP-1999; 99US-0159584P.
PR 27-SEP-1999; 99US-0160741P.
PR 28-SEP-1999; 99US-0160747P.
PR 29-SEP-1999; 99US-0160768P.
PR 30-SEP-1999; 99US-0160770P.
PR 01-OCT-1999; 99US-0160814P.
PR 02-OCT-1999; 99US-0160815P.
PR 03-OCT-1999; 99US-0160980P.
PR 04-OCT-1999; 99US-0160981P.
PR 05-OCT-1999; 99US-0160989P.
PR 06-OCT-1999; 99US-0161404P.
PR 07-OCT-1999; 99US-0161406P.
PR 08-OCT-1999; 99US-0161359P.
PR 09-OCT-1999; 99US-0161360P.
PR 10-OCT-1999; 99US-0161361P.
PR 11-OCT-1999; 99US-0161920P.

PR 28-OCT-1999; 99US-0161992P.
PR 29-OCT-1999; 99US-0161993P.
PR 30-OCT-1999; 99US-0162142P.

Query Match 54.3%; Score 44; DB 3; Length 292;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 EKTPLTTAAKAPVV 16
Db 73 EKAEWTTAMQSPVV 87

RESULT 7
ABB52865
ID ABB52865 standard; protein; 306 AA.
XX
AC ABB52865;
XX
DT 11-FEB-2002 (first entry)
XX
DE Escherichia coli polypeptide SEQ ID NO 1139.
XX
KW Escherichia coli; B2/D+A-; antiinflammatory; antibacterial;
immunosuppressive; extra-intestinal infection; phylogeny; meningitis;
systemic infection; non-diarrhoeal infection; septicaemia;
pyelonephritis; antibiotic resistance.
XX
OS Escherichia coli.
XX
PN WO200166572-A2.
XX
PD 13-SEP-2001.
XX
PF 12-MAR-2001; 2001WO-EP003445.
XX
PR 10-MAR-2000; 2000FR-00003145.
PR 02-FEB-2001; 2001FR-00001449.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Bingen E, Bonaccorsi S, Clermont O, Nassif X, Tinsley C;
WPI; 2001-550253/61.
XX
PT A library of DNA fragments of Escherichia coli strains for the phylogenetic
determination of a given strain comprises polynucleotides of nature B2/D+
A-.
XX
PS Example 6; Fig 6; 646pp; English.
XX
CC The invention relates to a library of DNA fragments of Escherichia coli
strains comprising polynucleotides (ABA8577-ABA88729 and ABA89533) and
encoded proteins (ABB52459-ABB52919 and ABB52954-ABB53094) of nature
B2/D+A-. The polynucleotides have potential antiinflammatory,
antibacterial and immunosuppressive activity as part of pharmaceutical
compositions used to treat, palliate or prevent extra-intestinal E. coli
infections. The polypeptides are useful for determining the phylogenetic
group of a given E. coli strain. These polypeptides can detect and treat
an undesired development of E. coli, particularly an extra-intestinal
infection that include systemic and non-diarrhoeal infections such as
septicaemia, pyelonephritis and meningitis this is particularly
advantageous as bacterial resistance is increasing with the more frequent
use of broad spectrum antibiotics
XX
SQ Sequence 306 AA;

Query Match 54.3%; Score 44; DB 4; Length 306;
Best Local Similarity 47.1%; Pred. No. 32;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 3 EKTPLTTAAKAPVVXNA 19
|||:::|:|

```

Db 84 EKVPCTSSGVAPIFVNA 100

RESULT 8
 AAB67588 AAB67588 standard; protein; 306 AA.
 XX AC AAB67588;
 XX DT 29-MAY-2001 (first entry)
 XX DE Amino acid sequence of a deoxyribokinase enzyme.
 XX KW Deoxyribonucleoside; deoxyribose 1-phosphate; thymidine phosphorylase;
 KW purine nucleoside phosphorylase; phosphopentose mutase;
 KW phosphopentose aldolase; fructose 1,6-diphosphate aldolase;
 KW deoxyribokinase; nucleoside 2-deoxyribosyltransferase.
 XX OS Salmonella typhi.
 XX PN WO200114566-A2.
 XX PD 01-MAR-2001.
 XX PF 18-AUG-2000; 2000WO-EP008088.
 XX PR 20-AUG-1999; 99EP-00116425.
 XX PA (HOFF) ROCHE DIAGNOSTICS GMBH.
 XX PA (INSP) INST PASTEUR.
 XX PA (PHAR-) PHARMA-WALDHOF GMBH & CO KG.
 XX PI Tischer W, Ihlenfeldt H, Barzu O, Sakamoto H, Pistonnik E;
 PI Marliere P, Pochet S;
 XX DR N-PSDB; AAF55444.
 XX PT In vitro enzymatic synthesis of deoxyribonucleosides comprises reacting
 PT deoxyribose 1-phosphate and a nucleobase to form a deoxyribonucleoside
 PT and an inorganic phosphate.
 XX PS Disclosure; Page 59-61; 73pp; English.
 XX CC The present sequence represents a deoxyribokinase enzyme. This enzyme is
 CC involved in the biosynthesis of deoxyribonucleosides, and is used in the
 CC method of the invention. The specification describes a method for the in
 CC vitro enzymatic synthesis of deoxyribonucleosides. The method comprises
 CC reacting deoxyribose 1-phosphate and a nucleobase to form a
 CC deoxyribonucleoside and an inorganic phosphate. Enzymes which may be used
 CC in the method of the invention include thymidine phosphorylase, purine
 CC nucleoside phosphorylase, phosphopentose mutase, phosphopentose aldolase,
 CC fructose 1,6-diphosphate aldolase, deoxyribokinase, and nucleoside 2-
 CC deoxyribosyltransferase
 XX SQ Sequence 306 AA;
 Query Match 54.3%; Score 44; DB 4; Length 306;
 Best Local Similarity 47.1%; Pred No. 32;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 3 EKPLTTTAAAPVWNA 19
 |||::|:
 Db 84 EKVPCTSSGVAPIFVNA 100

RESULT 9
 AAG30089 AAG30089 standard; protein; 427 AA.
 XX AC AAG30089;
 XX DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 35910.
 DE Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-00301439.
 XX PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.

```

PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140333P.
PR 24-JUN-1999; 99US-0140354P.
PR 25-JUN-1999; 99US-0140695P.
PR 26-JUN-1999; 99US-0140823P.
PR 27-JUN-1999; 99US-0140991P.
PR 28-JUN-1999; 99US-0141287P.
PR 29-JUN-1999; 99US-0141842P.
PR 30-JUN-1999; 99US-0142154P.
PR 01-JUL-1999; 99US-0142055P.
PR 02-JUL-1999; 99US-0142390P.
PR 03-JUL-1999; 99US-0142803P.
PR 04-JUL-1999; 99US-0142920P.
PR 05-JUL-1999; 99US-0142977P.
PR 06-JUL-1999; 99US-0143542P.
PR 07-JUL-1999; 99US-0143624P.
PR 08-JUL-1999; 99US-0144005P.
PR 09-JUL-1999; 99US-0144085P.
PR 10-JUL-1999; 99US-0144086P.
PR 11-JUL-1999; 99US-0144325P.
PR 12-JUL-1999; 99US-0144331P.
PR 13-JUL-1999; 99US-0144332P.
PR 14-JUL-1999; 99US-0144333P.
PR 15-JUL-1999; 99US-0144334P.
PR 16-JUL-1999; 99US-0144335P.
PR 17-JUL-1999; 99US-0144352P.
PR 18-JUL-1999; 99US-0144632P.
PR 19-JUL-1999; 99US-0144884P.
PR 20-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145088P.
PR 23-JUL-1999; 99US-0145085P.
PR 24-JUL-1999; 99US-0145087P.
PR 25-JUL-1999; 99US-0145089P.
PR 26-JUL-1999; 99US-0145192P.
PR 27-JUL-1999; 99US-0145145P.
PR 28-JUL-1999; 99US-0145218P.
PR 29-JUL-1999; 99US-0145224P.
PR 30-JUL-1999; 99US-0145276P.
PR 31-JUL-1999; 99US-0145913P.
PR 01-AUG-1999; 99US-0145918P.
PR 02-AUG-1999; 99US-0145919P.
PR 03-AUG-1999; 99US-0146386P.
PR 04-AUG-1999; 99US-0146388P.
PR 05-AUG-1999; 99US-0146389P.
PR 06-AUG-1999; 99US-0147038P.
PR 07-AUG-1999; 99US-0147204P.
PR 08-AUG-1999; 99US-0147302P.
PR 09-AUG-1999; 99US-0147192P.
PR 10-AUG-1999; 99US-0147260P.
PR 11-AUG-1999; 99US-0147303P.
PR 12-AUG-1999; 99US-0147416P.
PR 13-AUG-1999; 99US-0147493P.
PR 14-AUG-1999; 99US-0148171P.
PR 15-AUG-1999; 99US-0148319P.
PR 16-AUG-1999; 99US-0148341P.
PR 17-AUG-1999; 99US-0148565P.
PR 18-AUG-1999; 99US-0148684P.
PR 19-AUG-1999; 99US-0149175P.
PR 20-AUG-1999; 99US-0149426P.
PR 21-AUG-1999; 99US-0149722P.
PR 22-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 24-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 28-AUG-1999; 99US-0151066P.

PR 27-AUG-1999; 99US-0151080P.
PR 28-AUG-1999; 99US-0151303P.
PR 29-AUG-1999; 99US-0151438P.
PR 30-AUG-1999; 99US-0151930P.
PR 31-AUG-1999; 99US-0152363P.
PR 01-SEP-1999; 99US-0153070P.
PR 02-SEP-1999; 99US-0153758P.
PR 03-SEP-1999; 99US-0154018P.
PR 04-SEP-1999; 99US-0154039P.
PR 05-SEP-1999; 99US-0154779P.
PR 06-SEP-1999; 99US-0155139P.
PR 07-SEP-1999; 99US-0155486P.
PR 08-SEP-1999; 99US-0156559P.
PR 09-SEP-1999; 99US-0156558P.
PR 10-SEP-1999; 99US-0156596P.
PR 11-SEP-1999; 99US-0157117P.
PR 12-SEP-1999; 99US-0157753P.
PR 13-SEP-1999; 99US-0157865P.
PR 14-SEP-1999; 99US-0158029P.
PR 15-SEP-1999; 99US-0158232P.
PR 16-SEP-1999; 99US-0158369P.
PR 17-SEP-1999; 99US-0159293P.
PR 18-SEP-1999; 99US-0159294P.
PR 19-SEP-1999; 99US-0159295P.
PR 20-SEP-1999; 99US-0159330P.
PR 21-SEP-1999; 99US-0159331P.
PR 22-SEP-1999; 99US-0159637P.
PR 23-SEP-1999; 99US-0159638P.
PR 24-SEP-1999; 99US-0159584P.
PR 25-SEP-1999; 99US-0160741P.
PR 26-SEP-1999; 99US-0160767P.
PR 27-SEP-1999; 99US-0160768P.
PR 28-SEP-1999; 99US-0160770P.
PR 29-SEP-1999; 99US-0160814P.
PR 30-SEP-1999; 99US-0160815P.
PR 01-OCT-1999; 99US-0160815P.
PR 02-OCT-1999; 99US-0160980P.
PR 03-OCT-1999; 99US-0160981P.
PR 04-OCT-1999; 99US-0160989P.
PR 05-OCT-1999; 99US-0161404P.
PR 06-OCT-1999; 99US-0161405P.
PR 07-OCT-1999; 99US-0161406P.
PR 08-OCT-1999; 99US-0161359P.
PR 09-OCT-1999; 99US-0161360P.
PR 10-OCT-1999; 99US-0161361P.
PR 11-OCT-1999; 99US-0161920P.
PR 12-OCT-1999; 99US-0161922P.
PR 13-OCT-1999; 99US-0161993P.
PR 14-OCT-1999; 99US-0162142P.

Query Match 54.38; Score 44; DB 3; Length 427;
Best Local Similarity 60.0%; Pred. NO. 46;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 BEKTEPLTAAKAPV 16
Db 208 BEKTEPLTAAKAPV 222

RESULT 10
ABU19765
ID ABU19765 standard; protein; 584 AA.
AC ABU19765;
DT 19-JUN-2003 (first entry)
DE Protein encoded by prokaryotic essential gene #5292.
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS Borrelia cepacia.
XX

```

PN W0200277183-A2.
 XX 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WFI; 2003-029926/02.
 DR N-PSDB; ACA23635.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 47699; 1765pp; English.
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide; or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 584 AA;
 Query Match 54.3%; Score 44; DB 6; Length 584;
 Best Local Similarity 56.2%; Pred. No. 65;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 4 KPTPLTAAAXPVVXNA 19
 Db 101 RTPLSSTAAQVVVXNA 116
 RESULT 11
 AAG43604
 ID AAG43604 standard; protein; 172 AA.
 XX

AC AAG43604;
 XX 18-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 54519.
 DE Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX EP1033405-A2.
 XX 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-00301439.
 XX 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123549P.
 PR 23-MAR-1999; 99US-0125789P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 23-APR-1999; 99US-0130894P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 16-MAY-1999; 99US-0134370P.
 PR 19-MAY-1999; 99US-0134769P.
 PR 20-MAY-1999; 99US-0134941P.
 PR 21-MAY-1999; 99US-0135124P.
 PR 24-MAY-1999; 99US-0135353P.
 PR 25-MAY-1999; 99US-0135629P.
 PR 27-MAY-1999; 99US-0136021P.
 PR 28-MAY-1999; 99US-0136392P.
 PR 01-JUN-1999; 99US-0136782P.
 PR 03-JUN-1999; 99US-0137222P.
 PR 04-JUN-1999; 99US-0137528P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 16-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.

```
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 22-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144065P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.

PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154035P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0156659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161932P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.98; Score 42; DB 3; Length 172;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 2 EEXTPLTTAAAXAPVYXNA 19
    |||||
Db 131 EEXTFVEKGTGVPVYKKA 148

RESULT 12
AAG43603
ID AAG43603 standard; protein; 215_AA.
XX
AC AAG43603;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54518.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
```


KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130831P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135639P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139453P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 21-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 23-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 22-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.

```

PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161923P.
PR 28-OCT-1999; 99US-0161932P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.9%; Score 42; DB 3; Length 215;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 EEXTPLTTAAAXAPVWXA 19
Db 174 EEXTPEKKTGVPVWXA 191

RESULT 13
ID AAG43602 standard; protein; 332 AA.
XX
AC AAG43602;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54517.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX

```

```

PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140895P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.

```

```
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-AUG-1999; 99US-0145951P.
PR 28-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147303P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.

PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160800P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 51.9%; Score 42; DB 3; Length 332;
Best Local Similarity 50.0%; Pred. No. 76;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 EEKPLTAAAPVXNA 19
    |||||
Db 291 EEKTPVEKKGVPVVKKA 308

RESULT 14
ABUS8158
ID ABUS8158 standard; protein; 132 AA.
AC ABUS8158;
XX
XX 14-APR-2003 (first entry)
XX
XX Rice stress response protein #4.
DE
XX
XX Plant; EST; expressed sequence tag; stress response; drought; heat;
XX radiation; pathogen attack; grain flavour; disease resistance;
XX peptide-methionine sulphoxide reductase; DNA repair; enzyme;
XX intracellular protein transport.
XX
XX Oryza sativa.
XX
XX US2002152497-A1.
XX
XX 17-OCT-2002.
XX
XX 19-FEB-2002; 2002US-00078929.
XX
XX 07-MAY-1999; 99US-0133038P.
XX 07-MAY-1999; 99US-0133042P.
PR
```

PR 11-MAY-1999; 99US-0133427P.
 PR 11-MAY-1999; 99US-0133428P.
 PR 11-MAY-1999; 99US-0133436P.
 PR 11-MAY-1999; 99US-0133437P.
 PR 11-MAY-1999; 99US-0133438P.
 PR 04-JUN-1999; 99US-0137667P.
 PR 05-MAY-2000; 2000US-00566394.
 XX (FALC/) FALCO S. C.
 PA (FAMO/) FAMODU O. O.
 PA (MEYE/) MEYERS B. C.
 PA (MIAO/) MIAO G.
 PA (ODEL/) ODELL J. T.
 PA (RAFA/) RAFALSKI J. A.
 PA (THOR/) THORPE C. J.
 PA (SAKA/) SAKAI H.
 PA (WENG/) WENG Z.
 XX Falco SC, Famodu OO, Meyers BC, Miao G, Odell JT, Rafalski JA;
 PI Thorpe CJ, Sakai H, Weng Z;
 XX WPI; 2003-198391/19.
 DR N-PSDB; ABX78329.
 XX New peptide-methionine sulfoxide reductase and nucleic acids, useful in
 PT improving plant response to stress, engineering plants with increased
 PT disease and stress resistance, or and improving/protecting grain flavor.
 XX Claim 1; Fig 2; 205pp; English.
 XX The invention relates to isolated nucleic acids encoding plant stress
 CC response proteins (including peptide-methionine sulfoxide reductases)
 CC appearing as ABUS8148-ABUS8246 (or a protein 80% identical to them) from
 CC Zea mays, Oryza sativa, Glycine max, or Triticum aestivum. Also included
 CC are expression cassettes, transformed host cells, transgenic
 CC plants/seeds, modulating the level of peptide-methionine sulfoxide
 CC reductase in a plant and a computer system/data processing system for
 CC identifying, analysing, or modelling a genetic sequence. The plant
 CC nucleic acid is useful in developing strategies to improve plant response
 CC to stress (e.g. drought, heat, radiation or pathogen attack), engineering
 CC plants with increased disease and stress resistance, manipulating DNA
 CC repair and recombination efficiency, manipulating intracellular protein
 CC transport, and improving/protecting grain flavour. The nucleic acids may
 CC also be used as probes or amplification primers in the detection,
 CC quantitation or isolation of gene transcripts, for recombinant expression
 CC of encoded polypeptides, as immunogens in preparing or screening
 CC antibodies, and in sense or antisense suppression of one or more genes in
 CC a host cell, tissue or plant. The proteins may be used as immunogens or
 CC antigens to obtain antibodies specifically immunoreactive with the
 CC protein, and in assays for enzyme agonists or antagonists. The present
 CC sequence is a plant stress response protein (or fragment)
 XX Sequence 132 AA;
 SQ
 Query Match 50.6%; Score 41; DB 6; Length 132;
 Best Local Similarity 50.0%; Pred. No. 40;
 Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
 QY 2 EEXTPLTTAAAXAPVVXNA 19
 DB 97 EAAPPTTTAAEAFAIAAA 114
 RESULT 15
 ABU23415
 ID ABU23415 standard; protein; 517 AA.
 XX ABU23415;
 AC ABU23415;
 XX 19-JUN-2003 (first entry)
 DT Protein encoded by Prokaryotic essential gene #8942.
 DE
 XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX Bordetella pertussis.
 OS WO200277183-A2.
 PN 03-OCT-2002.
 PD 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 XX 06-SEP-2001; 2001US-00948993.
 PR 23-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA27285.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 25; SEQ ID NO 51339; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pt_sequences
 XX Sequence 517 AA;
 SQ
 Query Match 50.6%; Score 41; DB 6; Length 517;
 Best Local Similarity 75.0%; Pred. No. 1.8e+02;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 7 LTTAAAXAPVVXN 18
 DB 44 LTTAAGAPVADN 55

PN WO200170955-A2.
 XX 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haeselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS53967.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Example 3; SEQ ID NO 11701; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence represents an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 445 AA;
 XX
 Query Match 49.4%; Score 40; DB 4; Length 445;
 Best Local Similarity 57.1%; Pred. No. 2.3e+02;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 2 EEXTPLTTAAXAPV 15
 Db 382 EQKPLTTAADMMAI 395
 XX
 RESULT 19
 AAR51691
 ID AAR51691 standard; protein; 498 AA.
 XX
 AC AAR51691;
 XX
 XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 07-NOV-1994 (first entry)
 XX
 XX HIV-type virus WVP5180/91 gag protein (cloned).
 DE
 XX Human immunodeficiency virus; HIV; antigen; detection; diagnosis;
 XX retrovirus; vaccine; lymphocyte; reverse transcriptase.
 KW
 XX Unidentified.
 OS

XX PF591914-A2.
 XX 13-APR-1994.
 XX
 XX 05-OCT-1993; 93EP-00116058.
 XX
 XX 06-OCT-1992; 92DE-04233646.
 PR 22-OCT-1992; 92DE-04235718.
 PR 30-DEC-1992; 92DE-04244541.
 PR 01-JUN-1993; 93DE-04318186.
 XX
 PA (BEHW) BEHRINGWERKE AG.
 XX
 PI Guertler IG, Eberle J, Brunn VA, Knapp S, Hauser H;
 PI WPI; 1994-120077/15.
 DR N-PSDB; AAQ58974.
 XX
 PT New HIV-type immune deficiency virus ECACC V 92092318 - and deriv. cDNA
 PT or antigens, useful for diagnosing retroviral infections and vaccines.
 XX
 PS Disclosure; Fig 7; 73pp; German.
 XX
 CC MYP-5180/91 has been isolated from peripheral lymphocytes of a patient,
 CC from the Cameroons, with immune deficiency. It grows in the same human
 CC cells as HIV-1; like HIV it produces an Mg-dependent reverse
 CC transcriptase (RT), although this is 3 to 7 kD smaller in Western blots
 CC than the enzyme from HIV-1 or -2. It is less reactive than HIV-1, but
 CC more reactive than HIV-2, against p24-specific antibodies and its gp41
 CC transmembrane protein reacts with antibodies in sera of African patients
 CC but not (or only weakly) with sera of German patients. The virus consists
 CC of RNA in a peg-shaped core made of p24 subunits surrounded by an outer
 CC core of p17 and then a glycoprotein envelope which, apart from host-cell
 CC derived lipids, comprises gp41 and envelope protein gp120 (which can bind
 CC to the CD4 receptor). Related viruses have at least 75% homology over the
 CC entire genome with max. differences for the various regions LTR and NEF
 CC 10%, POL 12%, GAG 14%, VIF 15% and ENV 22%. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated
 CC on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 498 AA;
 XX
 Query Match 49.4%; Score 40; DB 2; Length 498;
 Best Local Similarity 44.4%; Pred. No. 2.6e+02;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 2 EEXTPLTTAAXAPVXNA 19
 Db 119 EETSPTTSQNYPIVTNA 136
 XX
 RESULT 20
 AAW93076
 ID AAW93076 standard; protein; 498 AA.
 XX
 AC AAW93076;
 XX
 XX 20-MAR-2003 (revised)
 DT 19-MAY-1999 (first entry)
 XX
 XX HIV isolate 5180 gag protein.
 DE
 XX HIV-type retrovirus; MYP-5180/91; ECACC V 92092318; antigen; assay kit;
 KW detection; antibody; immune deficiency; vaccine.
 KW
 OS Human immunodeficiency virus.
 XX
 PN EP890642-A2.
 XX
 XX 13-JAN-1999.
 PD
 XX 05-OCT-1993; 98EP-00114623.
 PF

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX
XX Disclosure; SEQ ID NO 8541; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 788 AA;
XX
XX Query Match 49.4%; Score 40; DB 4; Length 788;
XX Best Local Similarity 52.9%; Pred. No. 4.3e+02;
XX Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
XX
XX QY 3 EKTPLTTAAAPVXNA 19
XX | | | | | | | | | |
XX 602 EDGPTTTAAAPLASAA 618
XX
XX RESULT 23
XX ID AAE38171 standard; protein; 788 AA.
XX AC AAE38171;
XX AC AAE38171;
XX DT 20-NOV-2003 (first entry)
XX
XX Fruit fly G protein-coupled receptor (GPCR) protein #20.
XX
XX Fruit fly; G protein-coupled receptor; GPCR; cell therapy; nematode;
XX insect infestation; insecticide; nematocide; receptor.
XX
XX Drosophila melanogaster.
XX
XX WO2003052078-A2.
XX
XX 26-JUN-2003.
XX
XX 18-DEC-2002; 2002WO-US040525.
XX
XX 18-DEC-2001; 2001US-0341512P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Spana E, Kandar K, Stam L, Valentine S, Griswold CM;
XX WPI; 2003-533019/50.
XX N-PSDB; AAD57448.
XX
XX New polypeptide, useful for preventing or abrogating insect or nematode
XX infestation of a plant.
XX
XX Claim 1; Page 185-188; 368pp; English.
XX
XX The invention relates to a novel G protein-coupled receptor (GPCR)
XX polypeptide and its polynucleotide. The polypeptide is useful for
XX preventing or abrogating insect or nematode infestation of a plant. GPCR
XX polypeptide is useful in cell therapy. The present sequence is fruit fly
XX GPCR protein
XX
XX Sequence 788 AA;
XX
XX Query Match 49.4%; Score 40; DB 7; Length 788;

Best Local Similarity 52.9%; Pred. No. 4.3e+02;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 3 EKTPLTTAAAPVXNA 19
| | | | | | | | | |
DB 602 EDGPTTTAAAPLASAA 618

RESULT 24
ID ABB68024 standard; protein; 1473 AA.
XX ABB68024;
XX AC ABB68024;
XX DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 30864.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL12127.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 30864; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1473 AA;
XX
XX Query Match 49.4%; Score 40; DB 4; Length 1473;
XX Best Local Similarity 53.3%; Pred. No. 8.7e+02;
XX Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
XX
XX QY 2 EKTPLTTAAAPV 16
XX | | | | | | | | | |
XX 1315 DEQAPYTEALGPV 1329
XX
XX RESULT 25
XX ID ABB59613 standard; protein; 1793 AA.
XX AC ABB59613;
XX AC ABB59613;
XX


```

DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 5631.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW Pharmaceutical.
XX
XX Drosophila melanogaster.
OS WO200171042-A2.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
PF
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX
XX 11-JUL-2000; 2000US-00614150.
PR
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-656860/75.
DR
XX
XX N-PSDB; ABL03716.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
PT
XX
XX Disclosure; SEQ ID NO 5631; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1793 AA;
SQ
    Query Match      48.8%; Score 39.5; DB 4; Length 1793;
    Best Local Similarity 52.6%; Pred No. 1.3e+03;
    Matches 10; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

OY      2 EKKTLTT-AAXAPVXNA 19
Db      ::::|::|::|::|::|
        710 KKSSPTTTPARAPVAQNA 728

RESULT 26
AAG39690
ID AAG39690 standard; protein; 128 AA.
XX
XX AAG39690;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 49147.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX

```

```

PF 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR
XX
XX 05-MAR-1999; 99US-0123180P.
PR
XX
XX 09-MAR-1999; 99US-0123548P.
PR
XX
XX 23-MAR-1999; 99US-0125788P.
PR
XX
XX 25-MAR-1999; 99US-0126264P.
PR
XX
XX 29-MAR-1999; 99US-0126785P.
PR
XX
XX 01-APR-1999; 99US-0127462P.
PR
XX
XX 06-APR-1999; 99US-0128234P.
PR
XX
XX 16-APR-1999; 99US-0128714P.
PR
XX
XX 19-APR-1999; 99US-0130077P.
PR
XX
XX 21-APR-1999; 99US-0130449P.
PR
XX
XX 23-APR-1999; 99US-0130510P.
PR
XX
XX 28-APR-1999; 99US-0130891P.
PR
XX
XX 30-APR-1999; 99US-0131449P.
PR
XX
XX 04-MAY-1999; 99US-0132048P.
PR
XX
XX 05-MAY-1999; 99US-0132484P.
PR
XX
XX 06-MAY-1999; 99US-0132485P.
PR
XX
XX 07-MAY-1999; 99US-0132486P.
PR
XX
XX 11-MAY-1999; 99US-0132863P.
PR
XX
XX 14-MAY-1999; 99US-0134218P.
PR
XX
XX 14-MAY-1999; 99US-0134219P.
PR
XX
XX 14-MAY-1999; 99US-0134221P.
PR
XX
XX 18-MAY-1999; 99US-0134370P.
PR
XX
XX 19-MAY-1999; 99US-0134768P.
PR
XX
XX 20-MAY-1999; 99US-0134941P.
PR
XX
XX 21-MAY-1999; 99US-0135124P.
PR
XX
XX 24-MAY-1999; 99US-0135353P.
PR
XX
XX 25-MAY-1999; 99US-0135629P.
PR
XX
XX 27-MAY-1999; 99US-0136021P.
PR
XX
XX 28-MAY-1999; 99US-0136782P.
PR
XX
XX 01-JUN-1999; 99US-0137222P.
PR
XX
XX 04-JUN-1999; 99US-0137528P.
PR
XX
XX 07-JUN-1999; 99US-0137724P.
PR
XX
XX 08-JUN-1999; 99US-0138094P.
PR
XX
XX 10-JUN-1999; 99US-0138540P.
PR
XX
XX 10-JUN-1999; 99US-0138847P.
PR
XX
XX 14-JUN-1999; 99US-0139119P.
PR
XX
XX 16-JUN-1999; 99US-0139452P.
PR
XX
XX 16-JUN-1999; 99US-0139453P.
PR
XX
XX 17-JUN-1999; 99US-0139492P.
PR
XX
XX 18-JUN-1999; 99US-0139454P.
PR
XX
XX 18-JUN-1999; 99US-0139455P.
PR
XX
XX 18-JUN-1999; 99US-0139456P.
PR
XX
XX 18-JUN-1999; 99US-0139457P.
PR
XX
XX 18-JUN-1999; 99US-0139458P.
PR
XX
XX 18-JUN-1999; 99US-0139459P.
PR
XX
XX 18-JUN-1999; 99US-0139460P.
PR
XX
XX 18-JUN-1999; 99US-0139461P.
PR
XX
XX 18-JUN-1999; 99US-0139462P.
PR
XX
XX 18-JUN-1999; 99US-0139463P.
PR
XX
XX 18-JUN-1999; 99US-0139750P.
PR
XX
XX 21-JUN-1999; 99US-0139763P.
PR
XX
XX 22-JUN-1999; 99US-0139817P.
PR
XX
XX 22-JUN-1999; 99US-0139899P.
PR
XX
XX 23-JUN-1999; 99US-0140353P.
PR
XX
XX 23-JUN-1999; 99US-0140354P.
PR
XX
XX 24-JUN-1999; 99US-0140695P.
PR
XX
XX 28-JUN-1999; 99US-0140823P.
PR
XX
XX 29-JUN-1999; 99US-0140931P.
PR
XX
XX 30-JUN-1999; 99US-0141287P.
PR
XX
XX 01-JUL-1999; 99US-0141842P.
PR
XX
XX 01-JUL-1999; 99US-0142154P.
PR
XX
XX 02-JUL-1999; 99US-0142055P.
PR
XX
XX 06-JUL-1999; 99US-0142390P.
PR
XX
XX 08-JUL-1999; 99US-0142803P.
PR

```


XX WPI: 2001-639362/73.
 DR N-PSDB; AAS90526.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PT
 PS Claim 20; SEQ ID NO 56998; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 137 AA;
 Query Match 48.1%; Score 39; DB 4; Length 137;
 Best Local Similarity 63.6%; Pred. NO. 92;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 5 TPLTTAAXAPV 15
 Db : ||||| :
 104 TPLTTASATPL 114
 RESULT 28
 ADA35852
 ID ADA35852 standard; protein; 145 AA.
 XX
 AC ADA35852;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Acinetobacter baumannii protein #3013.
 XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Bretton G, Bush D;
 XX
 XX WPI: 2003-576092/54.
 DR N-PSDB; ADA31726.
 DR

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 XX Example; SEQ ID NO 7139; 328pp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX
 SQ Sequence 145 AA;
 Query Match 48.1%; Score 39; DB 6; Length 145;
 Best Local Similarity 66.7%; Pred. NO. 98;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 4 KTELTTAAXAPV 15
 Db : ||||| :
 35 RVELTAAQAQPV 46
 RESULT 29
 ADB08554
 ID ADB08554 standard; protein; 154 AA.
 XX
 AC ADB08554;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Alloiococcus otitis antigenic protein SEQ ID NO:2494.
 XX
 KW Alloiococcus otitidis; antigenic protein; immunogenic; immunisation;
 KW gene therapy; Gram-positive bacterium; infection.
 XX
 OS Alloiococcus otitis.
 XX
 PN WO2003048304-A2.
 XX
 PD 12-JUN-2003.
 XX
 PF 25-NOV-2002; 2002WO-US036123.
 XX
 PR 29-NOV-2001; 2001US-0333777P.
 PR 18-NOV-2002; 2002US-0426742P.
 XX
 PA (AMHP) WYETH HOLDINGS CORP.
 XX
 PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
 XX
 DR WPI: 2003-505284/47.
 DR N-PSDB; ADB06553.
 XX
 XX New Alloiococcus otitidis polynucleotides and polypeptides, useful for
 PT treating and diagnosing diseases, drug screening assays and monitoring of
 PT effects during drug clinical trials.
 XX
 PS Claim 33; SEQ ID NO 2494; 1019pp; English.
 XX
 CC The present invention describes an isolated polynucleotide (I) of
 CC Alloiococcus otitidis genomic DNA, which encodes an antigenic protein.
 CC Alloiococcus otitidis is a Gram-positive bacterium. Also described: (1)
 CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
 CC expression vector comprising the novel isolated polynucleotide (I), its
 CC complement, degenerate variant or fragment; (3) a genetically engineered
 CC host cell, transfected, transformed or infected with the vector of (2);

CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC composition comprising the polypeptide, its complement, biological
 CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC expression vector; (6) a pharmaceutical composition comprising the
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)
 CC immunising against *Alloccoccus* otitidis by administering to a host the
 CC immunogenic composition; (9) detecting and/or identifying *Alloccoccus*
 CC otitidis in the biological sample; (10) a kit comprising a container
 CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the
 CC genetically engineered host cell under conditions suitable to produce the
 CC polypeptide from the culture. (I) can be used in gene therapy. The
 CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC invention can be used for treating and diagnosing diseases, drug
 CC screening assays and monitoring of effects during drug clinical trials.
 CC The polynucleotides are useful for expressing and detecting *Alloccoccus*
 CC otitidis. The present sequence represents an *Alloccoccus* otitidis
 CC antigen protein from the present invention.
 XX
 SQ Sequence 154 AA;

Query Match 48.1%; Score 39; DB 6; Length 154;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 KTPLTTAAAPV 15
 DB 96 KTPLTTATREV 107

RESULT 30
 ABG77349
 ID. ABG77349 standard; protein; 159 AA.

XX AC ABG77349;
 XX DT 05-NOV-2002 (first entry)
 XX DE Selected Interacting Domain (SID) polypeptide #160.
 XX KW Yeast; selected interacting domain; SID; antifungal; cancer; cytostatic;
 XX KW neuroprotective; Candida infection; gene therapy;
 XX KW neurodegenerative disease.

XX OS Saccharomyces cerevisiae.
 XX PN WO200259255-A2.
 XX PD 01-AUG-2002.

XX PF 25-JAN-2002; 2002WO-EP001350.

XX PR 26-JAN-2001; 2001US-0264577P.

XX PA (HYBR-) HYBRIGENICS.

XX PI Legrain P;

XX DR WPI; 2002-619165/66.

XX DR N-PSDB; ABS62963.

XX PT New complex between two interacting bait and prey Saccharomyces
 PT cerevisiae polypeptides, useful for preventing or treating Candida
 PT infection, cancer or neurodegenerative diseases in a mammal.

XX PS Claim 6; Page 127; 196pp; English.

XX CC The invention relates to a complex between two interacting Saccharomyces
 CC cerevisiae polypeptides, comprising two Selected Interacting Domain (SID)
 CC polypeptides as bait and prey proteins. A pharmaceutical composition
 CC comprising the complex is useful for preventing or treating Candida
 CC infection, cancer and neurodegenerative diseases in a human or animal,

CC preferably in a mammal. This sequence represents a SID polypeptide of the
 CC invention
 XX
 SQ Sequence 159 AA;

Query Match 48.1%; Score 39; DB 5; Length 159;
 Best Local Similarity 61.5%; Pred. No. 1.1e+02;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 BEKPTLTAAAP 14
 DB 14 BESTPTATAAP 26

RESULT 31
 ABU11251
 ID ABU11251 standard; protein; 159 AA.

XX AC ABU11251;

XX DT 10-DEC-2002 (first entry)

XX DE Yeast selected interacting domain protein SEQ ID NO: 636.

XX KW Yeast; protein-protein interaction; Selected Interacting Domain;
 XX KW SID (RTM); secretion yield; cancer; neurodegenerative disease; fungicide;
 XX KW cytostatic; neuroprotective.

XX OS Saccharomyces cerevisiae.

XX PN WO200266504-A2.

XX PD 29-AUG-2002.

XX PF 14-FEB-2002; 2002WO-EP002299.

XX PR 16-FEB-2001; 2001US-0269266P.

XX PA (HYBR-) HYBRIGENICS.

XX PI Legrain P;

XX DR WPI; 2002-674913/72.

XX DR N-PSDB; ABT11568.

XX PT New protein-protein complexes of Saccharomyces cerevisiae, useful in drug
 PT screening or development, for developing yeast strains with better
 PT secretion yield of protein, or in gene therapy (e.g. to treat Candida
 PT infection or cancer).

XX PS Claim 6; Page 262; 357pp; English.

XX CC The present invention relates to complexes between Saccharomyces
 CC cerevisiae Selected Interacting Domain (SID (RTM)) proteins and coding
 CC sequences. The protein complexes of *S. cerevisiae* are useful in drug
 CC development, in screening drugs or agents that modulate the interaction
 CC of proteins, for developing yeast strains with better secretion yield of
 CC protein, and in gene therapy. The protein complexes, polypeptides and
 CC polynucleotides are useful for preventing or treating Candida infection,
 CC cancer or neurodegenerative diseases in humans or animals. The present
 CC sequence is a protein of the invention

XX SQ Sequence 159 AA;

Query Match 48.1%; Score 39; DB 5; Length 159;
 Best Local Similarity 61.5%; Pred. No. 1.1e+02;
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 BEKPTLTAAAP 14
 DB 14 BESTPTATAAP 26

RESULT 32
AAAG91909
ID AAG91909 standard; protein; 233 AA.
XX
AC AAG91909;
XX
DT 26-SEP-2001 (first entry)
XX
DE C glutamicum protein fragment SEQ ID NO: 5663.
XX
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis.
XX
OS Corynebacterium glutamicum.
XX
PN EP1108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
XX
PR 07-APR-2000; 2000JP-00159162.
XX
PR 03-AUG-2000; 2000JP-00280988.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
PI
DR WPI: 2001-376931/40.
XX
DR N-PSDB; AAH67128.
XX
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Claim 17; SEQ ID NO 5663; 246pp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium, and identifying a homologue of a gene derived from
CC corynebacterium. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
SQ Sequence 233 AA;
Query Match 48.1%; Score 39; DB 4; Length 233;
Best Local Similarity 66.7%; Pred. NO. 1.7e-02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 5 TPLTTAAAPV 16
DB 35 TPTTASAPV 46
RESULT 33
AAE22281
ID AAE22281 standard; protein; 321 AA.
XX
AC AAE22281;
XX
DT 25-JUL-2002 (first entry)
XX
DE Murine SPAS-1 partial protein.
XX
XX Murine; protective immunity; cancer; SPAS-1; T cell; pharmaceutical;
KW prostate; breast; cervix; ovary; placenta; colon; brain; lung; kidney;
KW chronic lymphocytic leukaemia; germ cell; vaccine; cytostatic; TRAMP;
KW transgenic adenocarcinoma mouse prostate.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
PH Misc-difference 155..156
FT /note= "Encoded by AGCTAAGCA"
FT
FT Misc-difference 175..176
FT /note= "Encoded by GGATAAGGG"
FT
FT Misc-difference 177..178
FT /note= "Encoded by ACCTAAGCC"
FT
FT Misc-difference 201..202
FT /note= "Encoded by TGCTAGTCC"
FT
FT Misc-difference 204..205
FT /note= "Encoded by TTCTGACCT"
FT
FT Misc-difference 224..225
FT /note= "Encoded by ATTGACAC"
FT
FT Misc-difference 239..240
FT /note= "Encoded by GCATGAAG"
FT
FT Misc-difference 249..250
FT /note= "Encoded by GGTGAGGG"
FT
FT Misc-difference 297..298
FT /note= "Encoded by TTCTAACTA"
FT
FT Misc-difference 300..301
FT /note= "Encoded by TTCTGACCT"
FT
FT Misc-difference 310..311
FT /note= "Encoded by AAATAAATC"
FT
FT Misc-difference 321
FT /label= "Unknown"
FT /note= "Encoded by AA"
XX
XX WO200224739-A2.
XX
XX 28-MAR-2002.
XX
XX 13-SEP-2001; 2001WO-US028621.
XX
XX 21-SEP-2000; 2000US-0234472P.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Allison JP, Passo M, Shastri N;
XX
XX WPI: 2002-362424/39.
XX
XX N-PSDB; AAD35418.
XX
XX New SPAS-1 protein or antigen obtained from TRAMP-C2 tumor cells, useful
PT as vaccine for treating or inhibiting cancer in patient, e.g. prostate,
PT breast, cervix, ovary, placenta, colon, brain, lung, kidney or germ cell
PT cancer.
XX
XX Claim 2; Fig 1A; 107pp; English.
XX
XX The invention relates to compounds and methods for inducing protective
CC immunity against cancer. The compounds provided include polypeptides that
CC contain at least one immunogenic portion of one or more SPAS-1 protein
CC and DNA molecules encoding them or antigen obtained from transgenic
CC adenocarcinoma mouse prostate (TRAMP)-C2 tumour cells. The immunogenic
CC portion of the SPAS-1 human homologue polynucleotides sequence, SPAS-1
CC antibody or its antigen-binding fragment, the antigen-presenting cell,
CC the T cell population and the pharmaceutical compositions are useful for
CC inhibiting the development of a cancer in a patient, specifically
CC prostate, breast, cervix, ovary, placenta, colon, brain, lung, kidney,
CC chronic lymphocytic leukaemia or germ cell cancer. In particular, these
CC compounds are useful as vaccines for inducing protective immunity against
CC cancer. They are also useful for diagnosing cancer and monitoring cancer
CC progression. The present sequence is murine SPAS-1 partial protein
XX
XX Sequence 321 AA;

Query Match 48.1%; Score 39; DB 5; Length 321;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TPLTTAAXAPVV 16
 :|||||
 Db 60 SPTTTAATMPVV 71

RESULT 34

AAE22283
 ID AAE22283 standard; protein; 395 AA.

XX AC AAE22283;

XX DT 25-JUL-2002 (first entry)

XX DE Murine normal SPAS-1 protein.

XX KW Murine; protective immunity; cancer; SPAS-1; T cell; pharmaceutical;
 KW prostate; breast; cervix; ovary; placenta; colon; brain; lung; kidney;
 KW chronic lymphocytic leukaemia; germ cell; vaccine; cytostatic; TRAMP;
 KW transgenic adenocarcinoma mouse prostate.

XX OS Mus sp.

XX FN WO200224739-A2.

XX PD 28-MAR-2002.

XX PF 13-SEP-2001; 2001WO-US028621.

XX PR 21-SEP-2000; 2000US-0234472P.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Allison JP, Fasso M, Shastri N;

XX DR WPI; 2002-362424/39.

XX DR N-PSDB; AAD35420.

XX PT New SPAS-1 protein or antigen obtained from TRAMP-C2 tumor cells, useful
 PT as vaccine for treating or inhibiting cancer in patient, e.g. prostate,
 PT breast, cervix, ovary, placenta, colon, brain, lung, kidney or germ cell
 PT cancer.

XX PS Claim 2; Fig 1E; 107pp; English.

XX CC The invention relates to compounds and methods for inducing protective
 CC immunity against cancer. The compounds provided include polypeptides that
 CC contain at least one immunogenic portion of one or more SPAS-1 protein
 CC and DNA molecules encoding them or antigen obtained from transgenic
 CC adenocarcinoma mouse prostate (TRAMP)-C2 tumour cells. The immunogenic
 CC portion of the SPAS-1 human homologue polynucleotides sequence, SPAS-1
 CC antibody or its antigen-binding fragment, the antigen-presenting cell,
 CC the T cell population and the pharmaceutical compositions are useful for
 CC inhibiting the development of a cancer in a patient, specifically
 CC prostate, breast, cervix, ovary, placenta, colon, brain, lung, kidney,
 CC chronic lymphocytic leukaemia or germ cell cancer. In particular, these
 CC compounds are useful as vaccines for inducing protective immunity against
 CC cancer. They are also useful for diagnosing cancer and monitoring cancer
 CC progression. The present sequence is murine normal SPAS-1 protein

XX SQ Sequence 395 AA;

Query Match 48.1%; Score 39; DB 5; Length 395;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TPLTTAAXAPVV 16
 :|||||
 Db 300 SPTTTAATMPVV 311

RESULT 35

AAE22282
 ID AAE22282 standard; protein; 395 AA.

XX AC AAE22282;

XX DT 25-JUL-2002 (first entry)

XX DE Murine tumour SPAS-1 mutant protein.

XX KW Murine; protective immunity; cancer; SPAS-1; T cell; pharmaceutical;
 KW prostate; breast; cervix; ovary; placenta; colon; brain; lung; kidney;
 KW chronic lymphocytic leukaemia; germ cell; vaccine; cytostatic; TRAMP;
 KW transgenic adenocarcinoma mouse prostate; mutant; mutein.

XX OS Mus sp.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Region 244..252

XX FT /note= "Antigenic epitope capable of activating TRAMP-
 specific murine T cells"

XX FT Misc-difference 251

XX FT /note= "Wild-type Arg is substituted with His"

XX FN WO200224739-A2.

XX PD 28-MAR-2002.

XX PF 13-SEP-2001; 2001WO-US028621.

XX PR 21-SEP-2000; 2000US-0234472P.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Allison JP, Fasso M, Shastri N;

XX DR WPI; 2002-362424/39.

XX DR N-PSDB; AAD35419.

XX PT New SPAS-1 protein or antigen obtained from TRAMP-C2 tumor cells, useful
 PT as vaccine for treating or inhibiting cancer in patient, e.g. prostate,
 PT breast, cervix, ovary, placenta, colon, brain, lung, kidney or germ cell
 PT cancer.

XX PS Claim 2; Fig 1D; 107pp; English.

XX CC The invention relates to compounds and methods for inducing protective
 CC immunity against cancer. The compounds provided include polypeptides that
 CC contain at least one immunogenic portion of one or more SPAS-1 protein
 CC and DNA molecules encoding them or antigen obtained from transgenic
 CC adenocarcinoma mouse prostate (TRAMP)-C2 tumour cells. The immunogenic
 CC portion of the SPAS-1 human homologue polynucleotides sequence, SPAS-1
 CC antibody or its antigen-binding fragment, the antigen-presenting cell,
 CC the T cell population and the pharmaceutical compositions are useful for
 CC inhibiting the development of a cancer in a patient, specifically
 CC prostate, breast, cervix, ovary, placenta, colon, brain, lung, kidney,
 CC chronic lymphocytic leukaemia or germ cell cancer. In particular, these
 CC compounds are useful as vaccines for inducing protective immunity against
 CC cancer. They are also useful for diagnosing cancer and monitoring cancer
 CC progression. The present sequence is murine tumour SPAS-1 mutant protein

XX SQ Sequence 395 AA;

Query Match 48.1%; Score 39; DB 5; Length 395;
 Best Local Similarity 66.7%; Pred. No. 3e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 TPLTTAAXAPVV 16
 :|||||
 Db 300 SPTTTAATMPVV 311

RESULT 36

ABU38941
 ID ABU38941 standard; protein; 484 AA.
 AC ABU38941;
 XX
 XX 19-JUN-2003 (first entry)
 DT DT
 XX
 XX Protein encoded by Prokaryotic essential gene #24468.
 DE
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW
 XX Pasteurella multocida.
 OS
 XX WO200277183-A2.
 PN
 XX
 PD 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 PR
 XX 06-SEP-2001; 2001US-00948993.
 PR
 XX 25-OCT-2001; 2001US-0342323P.
 PR
 XX 08-FEB-2002; 2002US-00072851.
 PR
 XX 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACAM2811.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 6865; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *X. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 484 AA;

Query Match 48.1%; Score 39; DB 6; Length 484;
 Best Local Similarity 66.7%; Pred. No. 3.7e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 LTTAAAPVYVN 18
 |||||
 Db 13 LTNAGAPIVEN 24

RESULT 37

ADE31679

ID ADE31679 standard; protein; 2080 AA.

XX ADE31679;

AC ADE31679;

XX 29-JAN-2004 (first entry)

DT

XX Human 7077 protein #SEQ ID 36.

DE

XX

XX

KW Antiarteriosclerotic; cardiant; vasotropic; antiinflammatory;

KW thrombolytic; antiarrhythmic; antidiabetic; antihypertensive; gene therapy;

KW cardiovascular disorder; ischaemia; aortic bending;

KW vascular heart disease; endocarditis; atrial fibrillation; heart failure;

KW angina; cardiomyopathy; cardiac death.

XX

OS Homo sapiens.

XX

PN WO2003065984-A2.

XX

PD 14-AUG-2003.

XX

XX 29-JAN-2003; 2003WO-US002571.

XX

PR 01-FEB-2002; 2002US-0353224P.

PR 15-MAR-2002; 2002US-0364529P.

PR 19-APR-2002; 2002US-0373861P.

PR 29-APR-2002; 2002US-0376287P.

PR 12-JUN-2002; 2002US-0388080P.

PR 24-JUN-2002; 2002US-0390971P.

PR 03-JUL-2002; 2002US-0394130P.

PR 10-JUL-2002; 2002US-0394797P.

PR 21-AUG-2002; 2002US-0404904P.

PR 23-AUG-2002; 2002US-0405450P.

PR 04-SEP-2002; 2002US-0406070P.

PR 06-NOV-2002; 2002US-0424300P.

PR 05-DEC-2002; 2002US-0431042P.

PR 05-DEC-2002; 2002US-0431079P.

XX (MILL-) MILLENNIUM PHARM INC.

XX

PI Logan TJ, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;

PI Stagliano N, Perodin J, Rodrigue-Way A;

XX

XX WPI; 2003-731468/69.

DR N-PSDB; ADE31678.

XX

PT Identifying a compound capable of treating a cardiovascular disorder

PT (e.g. atherosclerosis) comprises assaying the ability of the compound to

PT modulate the expression or activity of e.g. 1682, 6169 or 6193

PT polypeptide or nucleic acid.

XX

PS Disclosure; SEQ ID NO 36; 328pp; English.

XX

CC The invention relates to a method for identifying a compound capable of

CC treating a cardiovascular disorder. The present invention identifies the

CC differential expression of 1682, 6169, 6193, 7771, 14395, 29002, 33216,

CC 43726, 69292, 21656, 32427, 2402, 7747, 1720, 9151, 60491, 1371, 7077,

CC 33207, 1419, 18036, 16105, 38650, 14245, 58848, 1870, 25856, 32394, 3484,

CC 345, 9252, 9135, 10532, 18610, 8165, 2448, 64624, 84237, 8912,

CC 2868, 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914, 51130,

CC 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474, 9792, 15400, 1452 or

CC

6585 genes in cardiovascular disease states. The methods are useful in diagnosing, preventing and treating cardiovascular disorders, such as atherosclerosis, cardiac hypertrophy, ischaemia, reperfusion injury, stenosis, arterial inflammation, vascular wall remodeling, coronary microembolism, tachycardia, bradycardia, pressure overload, aortic bending, coronary artery ligation, vascular heart disease, valvular disease, including but not limited to, valvular degeneration caused by calcification, rheumatic heart disease, endocarditis, or complications of artificial valves; atrial fibrillation, long-QT syndrome, congestive heart failure, sinus node dysfunction, angina, heart failure, hypertension, atrial fibrillation, atrial flutter, pericardial disease, including but not limited to, pericardial effusion and pericarditis; cardiomyopathies, e.g. dilated cardiomyopathy or idiopathic cardiomyopathy, myocardial infarction, coronary artery disease, coronary artery spasm, ischaemic disease, arrhythmia, sudden cardiac death and cardiovascular developmental disorders. The methods may also be used for identifying compounds that modulate cardiovascular disorders. Sequences given in ADE31644-ADE31769 represent the genes and proteins that may be regulated by a compound of the invention.

XX SQ Sequence 2080 AA;

Query Match 48.1%; Score 39; DB 7; Length 2080;
Best Local Similarity 47.1%; Pred. No. 1.9e+03;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 3 EKTPLTTAAXAPVVXNA 19
:||||:|
Db 135 QKTLGACIAPLKQGA 151

RESULT 28

ABE67966
ID ABE67966 standard; protein; 842 AA.

XX AC ABE67966;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 30690.

XX KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

XX KW Drosophila melanogaster.

XX OS WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL12069.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

XX PS Disclosure; SEQ ID NO 30690; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-CC ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 842 AA;

Query Match 47.5%; Score 38.5; DB 4; Length 842;
Best Local Similarity 50.0%; Pred. No. 8.4e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

OY 3 EKTPLTTAAXAPVVXNA 19
:||||:|
Db 598 EKTPIYSTRKAPVVST 615

RESULT 39

ADC32844
ID ADC32844 standard; protein; 203 AA.

XX AC ADC32844;

XX DT 18-DEC-2003 (first entry)

XX DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:2926.

XX KW Human; diagnostic; drug screening; forensics; gene mapping;

XX KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

XX KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

XX KW ulcers; osteoporosis; autoimmune disease; cancer;

XX KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

XX KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulvovaginal;

XX KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic; gene therapy; chromosome 11.

XX OS Homo sapiens.

XX PN WO2003029271-A2.

XX PD 10-APR-2003.

XX PF 24-SEP-2002; 2002WO-US030474.

XX PR 24-SEP-2001; 2001US-0324631P.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao CA, Wang J, Wehrman T;

XX PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

XX PI Haley-Vicente D, Drmanac RT;

XX DR WPI; 2003-371981/35.

XX DR N-PSDB; ADC32077.

XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or treating conditions such as neurodegenerative diseases, anemias, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer.

XX PS Example 2; SEQ ID NO 2926; 1185pp; English.

XX CC The invention relates to 971 novel human cDNA sequences (ADC29919-ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The invention also relates to nucleic acid sequences over 99% identical with the novel human cDNAs. The invention additionally encompasses expression vectors and host cells comprising a nucleic acid of the invention; the recombinant production of a polypeptide of the invention; an antibody against a polypeptide of the invention; a method of detecting polynucleotides or polypeptides of the invention; and methods of polynucleotides or polypeptides of the invention.

XX CC identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition; kits comprising polynucleotide probes

CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a human contig-
 CC encoded polypeptide sequence used in an example of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 203 AA;

Query Match 46.9%; Score 38; DB 7; Length 203;
 Best Local Similarity 46.7%; Pred. No. 2.1e-02;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 TPLTTAAAXAPVXNA 19
 ||: ||: ||: ||:
 Db 120 TTVPRANASPIVSGA 134

RESULT 40

ABB71511
 ID ABB71511 standard; protein; 225 AA.

AC ABB71511;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 41325.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL15614.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.

PS Disclosure; SEQ ID NO 41325; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 225 AA;

Query Match 46.9%; Score 38; DB 4; Length 225;
 Best Local Similarity 58.3%; Pred. No. 2.4e-02;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 TPLTTAAAXAPVV 16
 ||: ||: ||: ||:
 Db 201 TPTPTAPTAPVI 212

Search completed: October 4, 2004, 15:11:38
 Job time : 236 secs

B/SNK